WORKing papers in Management Science

WORMS/16/13

Eye tracking based experimental study on basic digital control panel usability

Rafał Michalski

Wrocław University of Science and Technology, Poland

WORMS is a joint initiative of the Management Science departments
of the Wrocław University of Science and Technology,
Wyb. Wyspiańskiego 27, 50-370 Wrocław, Poland

Published in:

Michalski R. (2016). Eye tracking based experimental study on basic digital control panel usability. [In:] European Network Intelligence Conference – ENIC 2016, 5-7 September, Wroclaw, Poland, IEEE, pp. 145-152. http://dx.doi.org/10.1109/ENIC.2016.029

Also available at: RafalMichalski.com

Eye Tracking Based Experimental Study on Basic Digital Control Panel Usability

Rafał Michalski

Faculty of Computer Science and Management, Wrocław University of Science and Technology, Poland

e-mail: rafal.michalski@pwr.edu.pl, www: RafalMichalski.com

***Abstract*—Various versions of a typical control panels are experimentally analyzed both from the effectiveness and efficiency point of view. The stimuli differed in objects' locations within a panel, and the type of the expected response to the presented stimuli. The results of users' performance are investigated using classical analyses of variance. Moreover, visual activity of all examined persons gathered by means of eye tracking system were presented and formally analyzed.**

***Keywords*—digital panel efficiency; digital panel effectiveness; visual activity; human-computer interaction; ergonomics**

I. Introduction

In various aspects of our life computers are more and more common and play more important roles. In many cases traditional solutions are replaced by their digital counterparts. During recent years, even in products like household appliances one may observe the tendency of using digital instead of classic control panels. Such ubiquity of digital control panels makes the usability aspects of their design especially important and worth studying.

There is a great number of various usability definitions provided by many authors, however one of the simplest and the most commonly used is the proposal described in ISO 9241-11 [15]. According to this international standard usability is characterized in three main dimensions, namely effectiveness, efficiency and satisfaction, taking into account the context of use.

Various types of control panels and usability aspects have been subject to investigations in previous papers. For instance, Tan et al. [26] examined the visibility of controls and labels in almost 100 versatile electronic appliances. Reference [19], among other things, provided a thorough review of control panels design guidelines for audiovisual consumer electronic products. The human-device usability analysis of manual defibrillators was subject to examination in [11]. The automatic air conditioning control panel was investigated by means of formal modeling in [7]. An innovative way of assessing interactions with household appliances' control panels involving virtual reality was described in [2].

The use of human visual activity data for usability purposes has a long history which was reviewed in [9], [16]. Generally, as it was stressed in many previous research (e.g. [8], [13]), eye tracking data may be helpful during the evaluation process because it provides information about how users allocate and shift their attention while interacting with a human-machine interface. Though, there are many investigations concerned with ergonomics and usability of conventional control panels and general ergonomic design principles, the body of literature on their digital equivalents (i.e. presented by software) where eye trackers are involved is not so ample. Among works in this field there is a usability evaluation of multi-modal interface for car drivers [22]. The authors developed a methodology for evaluating such interfaces that takes into account, among other things, the oculographic data. In a study of [1] on control panels researchers examined usability of portable ultrasound systems interfaces by means of the motion and eye balls movement analysis.

In the literature, panel layouts were most often analyzed and optimized from a perspective of moving a hand or mouse pointers between panel components (e.g. [14], [20], [21], [24]) The present paper examines the usability of specific control panel layout types by analyzing the number of errors made (effectiveness) and task execution times (efficiency). These basic analyses are deepened by the detailed exploration of human visual activity temporal parameters, registered during experimental tasks. Moreover, the conducted in this study analysis of pupil diameter changes may be interpreted to some extent as a satisfaction measure (e.g. [10], [23]). What is important, such analyses are conducted based on objective data which is quite rare in investigating this dimension.

II. Method

A. Participants

Overall, 50 student volunteers participated in the study, however the data of four persons with the eye tracking ratio smaller than 70% were excluded from visual activity analysis. The age ranged from 18 to 34 years with the average of 21.3 and a standard deviation of 3.5. There were 24 men and 22 women taking part in the study.

B. Apparatus

Simple, custom-made software was developed to present stimuli in a typical web browser window. The application was prepared in plain JavaScript and not only controlled the presentation of examined control panels but also registered task execution times and errors made. The experiment was carried out in identical lighting conditions in an isolated room equipped with a desk, typical office chair, keyboard, optical computer mouse, and 21 inches monitor with a resolution set at 1680×1050 pixels, with a classic color scheme. The subjects' behavior was monitored through a one way mirror installed in one of the room's walls. Communication with subjects took place via a set of microphones and speakers.

A SMI RED500 stationary infrared eye-tracker system was used to register oculographic activities while performing experimental tasks. The system comprised of the detector placed directly under the monitor displaying stimuli and the laptop computer working under *Microsoft Windows 7* operating system with *SMI Experiment Center 3.6* software by which the current experiment was controlled. The system recorded eye ball movements at 500 Hz sample rate. The *SMI BeGaze 3.6* application allowed for exporting raw fixation data for analyses performed in a *Statistica 12* package. The experimental application displaying stimuli was run on a *Mozilla Firefox* web browser ver. 45 set to the full screen mode.

C. *Stimuli, experimental design, dependent measures*

The current research focuses on operation of various variants of control panels. Digital versions of artificially created panels consisted generally of control buttons and information displays. The experimental setup resembled the situation when a car is being tested by a chassis dynamometer. In such a situation the operator may perform a monitoring task by comparing the current and target velocity and clicking on the appropriate button to achieve or maintain the desired test speed.

The experimental task was performed on simplified versions of possible virtual graphical interfaces. The prepared panels were differentiated by the way their components were located in relation to each other. There were three main panel sections, that is: the area with (1) control buttons (*Increase*, *Keep*, *Decrease*), (2) information about the current velocity presented on an analogue (clock-face) version of a speedometer, and, and finally (3) information about a target velocity presented by a digital display. In the remainder of the paper these components are denoted by *B*, *C*, and *T* characters respectively.

There are six different possibilities of arranging these sections: *BCT*, *BTC*, *CBT*, *CTB*, *TBC*, *TCB*, and they were treated as six levels of the *Layout* factor examined in the current study. The dimensions of components with buttons and information about target velocity were identical while the current velocity information section had twice as big area than the other sections.

Subjects taking part in the experiment performed nine trials per each layout type. These trials differed in two aspects, namely: the target velocity location (*TVL*) on the speedometer, which was specified on three levels (*20*, *50*, and *80* kilometers per hour), and the correct response type (*CRT*) involving also three levels (*Increase*, *Keep*, and *Decrease*). By combining the independent factor levels and trial variations one obtains: *Layout* (6) × *TVL* (3) × *CRT* (3) = 54 different experimental conditions. Some of them are demonstrated in Fig. 1.

Fig. 1. An exemplary stimuli presented by the experimental software in a web browser.

A within subjects design was used therefore each participant performed the experimental task for all prepared conditions that were presented by the experimental software. The presentation order was first determined randomly and then slightly modified in such a way that the same levels of the *Layout* effect would not appear one after another. Finally, the Latin square approach was applied to display stimuli to consecutive subjects.

The efficiency of the elaborated control panels operation was assessed based on the time necessary to make a decision, point and click with a computer mouse on the appropriate control button. The effectiveness was measured by the number of incorrect responses. Apart from the abovementioned dependent variables, fixation durations, fixation counts, and pupil diameters were additionally used to analyze subjects' visual activity during the experiments.

The fixations were determined by the *SMI BeGaze 3.6* software according to the High Speed Event Detection method with default parameters. This approach first searches for saccade events and then computes fixations from them. For more details about this type of saccades and fixations detection algorithm refer to [25] or BeGaze Manual [3].

D. *Experimental procedure*

Participants were first provided information about the goal and procedure of the experiment. The study began with a short questionnaire in which subjects were asked about their gender, age, visual acuity and potential vision disorders. Before performing main experimental tasks, they executed three test trials after which, the fast, two points eye tracker calibration procedure was applied. Each time between consecutive control panels, a white window with a black cross (x) at the center appeared for 1500 ms.

Participants were instructed to focus their eyes on this cross between trials and perform the task as quickly and accurately as possible. Subjects' visual activities were registered by the eye tracking system only while performing the proper experiment. The experimental trial sequence is illustrated in Fig. 2.

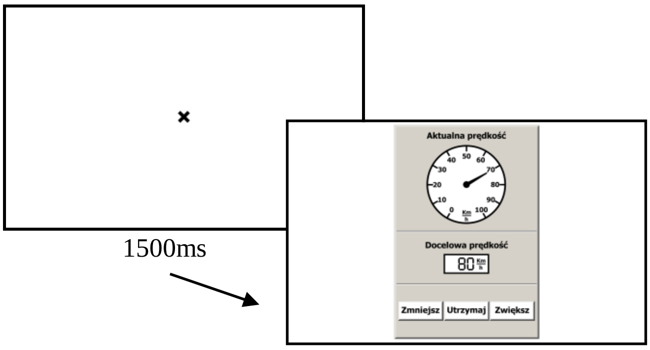

Fig. 2. A single trial sequence presented by experimental software.

III. Results

The results presented below concern three main areas related with usability, namely: task execution efficiency, and effectiveness as well as visual activity patterns registered by the eye tracking system. In all Analysis of Variance (Anova) figures presented in the following sections vertical bars denote 0.95 confidence intervals.

A. Task execution efficiency

The overall mean value of task performance for all participants and experimental conditions equaled 2094 ms, with the standard deviation 279 ms. The lowest average task execution time was observed for the condition with *TBC* panel layout, *50 km/h* target velocity, the *Keep* correct response type (*TBC_50_Keep*) and amounted to 1508 ms (Mean Standard Error, *MSE* = 105). The worst mean result 2484 (*MSE* = 105) was registered for *CBT* layout, *20 km/h*, and the *Decrease* correct response (*CBT_20_Decrease*). The task execution efficiency data were formally verified by means of the four-way Anova: (*Gender*) × (*Layout*) × (*Target velocity location*) × (*Correct response type)*. The analysis revealed that *Gender* and *Correct Response Type* factors significantly differentiated mean task execution times at the level $\alpha = 0.05$ ($F_{Gender}(1, 2592) = 8.24$, $p = 0.0041$, $\eta^2 = 0.0032$; $F_{CRT}(2, 2592) = 163$, $p < 0.0001$, $\eta^2 = 0.11$). The *Layout* and *Target Velocity Location* effects along with all two- three- and four-way interactions were irrelevant ($\alpha = 0.05$). The statistically significant results are illustrated in Figs. 3 and 4.

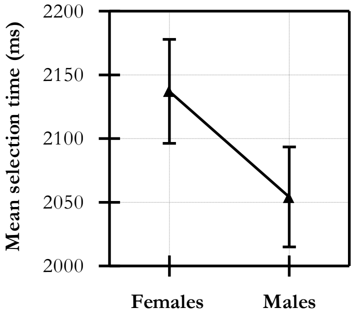

Fig. 3. *Gender* effect on mean task execution time. $F_{Gender}(1, 2592) = 8.24$, $p = 0.0041$,

They show significantly longer mean task execution times for women than for men (2137 ms ± 21 *MSE* vs. 2054 ms ± 20 *MSE*). Anova data regarding the *Correct response type* effect indicate that subjects performed tasks decidedly faster for the *Keep* option than for *Decrease* and *Increase*. As series of LSD Fisher's pairwise comparisons revealed the differences between *Keep* vs. *Decrease* and *Keep* vs. *Increase* were statistically relevant ($\alpha = 0.0001$) while mean tasks executions difference between *Decrease* vs. *Increase* was not significant ($\alpha = 0.1$).

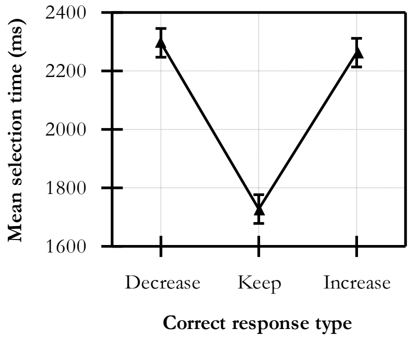

Fig. 4. *Correct response type* effect on mean task execution time. $F_{CRT}(2, 2592) = 163$, $p < .0001$.

B. Wrong selections

In total, subjects committed 150 errors which accounted for 5.5% of all performed trials. On average men made more mistakes (5.77%) than women (5.32%), however the difference was statistically irrelevant ($\chi^2 = 0.25$, $p = 0.62$). A Chi-squared test also showed no differences in the number of committed errors depending on the control panel layouts ($\chi^2 = 6.73$, $p = 0.24$) and *Target velocity location* ($\chi^2 = 4.7$, $p = 0.095$). Meaningful discrepancies were detected for the *Correct response type* ($\chi^2 = 87$, $p < 0.0001$). In this case subjects committed only 0.33% mistakes for the *Keep* type response, 6% for *Increase* and as much as 10.4% for the *Decrease* condition.

C. Visual activity

The gathered visual activity data were first analyzed for whole panels and a more in-depth insight was made by exploring oculographic parameters in particular parts of the presented stimuli, that is, sections containing control buttons, target and current velocity. In both cases appropriate Area of Interests (AOI) were defined in the BeGaze 3.6. application.

1) Analyses of variance for whole panels

A series of three standard three-way (*Gender*) × (*Layout*) × (*Correct response type)* Anovas were performed to test the significance of differences for three dependent variables: the fixation duration, fixation count, and pupil diameter. Due to the restriction of BeGaze 3.6. software in creating custom trial stimuli to 50, it was necessary to exclude from further analyses one of the experimental design factors. In order to minimize the data loss it was justified to group data according to one of the three levels factors. The *Target velocity location* factor was excluded as its impact on the user behavior (that is, number of incorrect responses, and mean task execution times) presented in previous sections of this paper was smaller than for the *Correct response type* effect.

a) Fixation durations

A three-way Anova was employed to check whether the *Gender (females, males)*, *Layout (BCT, BTC, CBT, CTB, TBC, TCB)*, and *Correct response type (Decrease, Keep, Increase)* factors have an effect on mean fixation durations. The obtained results suggest significant impact of all main factors: $F_{Gender}(1, 16\,786) = 55$, $p < 0.0001$, $\eta^2 = 0.0033$; $F_{CRT}(2, 16\,786) = 8.9$, $p = 0.00014$, $\eta^2 = 0.00106$; $F_{Layout}(5, 16\,786) = 2.7$, $p = 0.02$, $\eta^2 = 0.0008$. Additionally, the analysis revealed meaningful effect of (*Layout* × *CRT*) interaction on average fixation durations: $F_{Layout \times CRT}(10, 16\,786) = 1.9$, $p = 0.037$, $\eta^2 = 0.00115$. All other two and three way interactions were statistically insignificant ($\alpha = 0.05$). Mean fixation durations for meaningful effects are illustrated in Figs. 5-8. Fig. 5 demonstrates markedly longer mean fixation duration for women (148 ms ± 0.99 *MSE*) vs. men (137 ± 1.01 *MSE*).

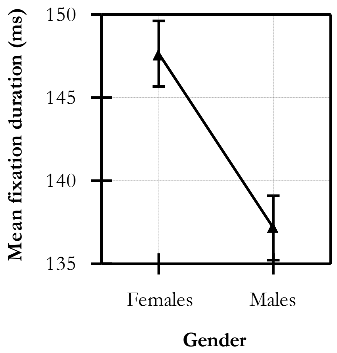

Fig. 5. *Gender* effect on mean fixation duration for whole panels. $F_{Gender}(1, 16\,786) = 55$, $p < 0.0001$, $\eta^2 = 0.0033$.

Data from Fig. 6 indicate that the shortest mean fixation duration was for the *Keep* response type (138 ± 1.35 *MSE*), distinctly longer was the average fixation duration for the *Increase* condition, (144 ± 1.15 *MSE*) whereas the longest one was detected for the *Decrease* case (145 ± 1.14 *MSE*).

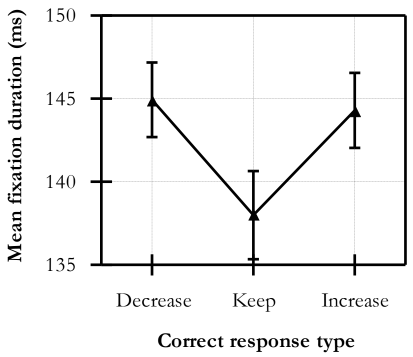

Fig. 6. *Correct response type* effect on mean fixation duration for whole panels. $F_{CRT}(2, 16\,786) = 8.9$, $p = 0.00014$, $\eta^2 = 0.00106$.

A series of LSD Fisher's analyses revealed that the difference in means for *Increase* and *Decrease* levels were irrelevant ($p = 0.69$) while other level pairs differed significantly ($\alpha = 0.0005$).

Results illustrated in Fig. 7 show that the shortest mean fixation duration was obtained for the *CBT* panel layout (138 ms ± 1.67 *MSE*) while the longest were computed for *CTB* condition (146 ms ± 1.71 *MSE*). LSD Fisher's post-hoc comparisons revealed significant differences between the following pairs of *Layout* versions: *CTB* vs. *BCT* ($p = 0.01$) and *CBT* ($p = 0.003$); *BTC* vs. *CBT* ($p = 0.02$); *CBT* vs. *TCB* ($p = 0.048$).

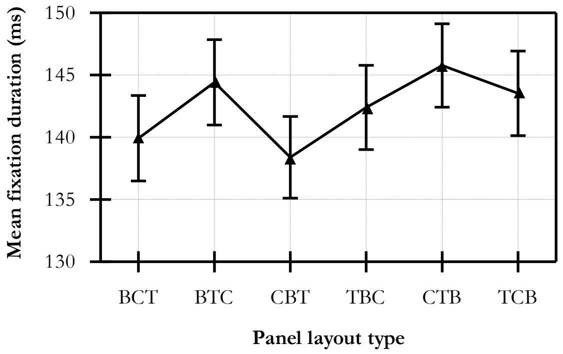

Fig. 7. *Panel layout type* effect on mean fixation duration for whole panels. $F_{Layout}(5, 16\,786) = 2.7$, $p = 0.02$, $\eta^2 = 0.0008$.

While analyzing fixation duration means for the *Layout* and *Target velocity location* interaction presented in Fig. 8 one may observe that generally values for the *Keep* and *Decrease* variants of *Correct response type* effect follow the pattern given in Fig. 7. For the *Keep* condition, however, fixation duration means are always smaller than for the *Decrease* option. As far as the *Increase* case is concerned, it may be noticed that for *Layouts BCT* and *BTC* the means are smaller in comparison with the *Decrease* condition, but for *CBT* and *TCB* the situation is reversed: longer average fixation durations are for the *Increase* condition. Finally, for the remaining *TCB* and *CTB* layouts the means are almost identical for the *Increase* and *Decrease* variants.

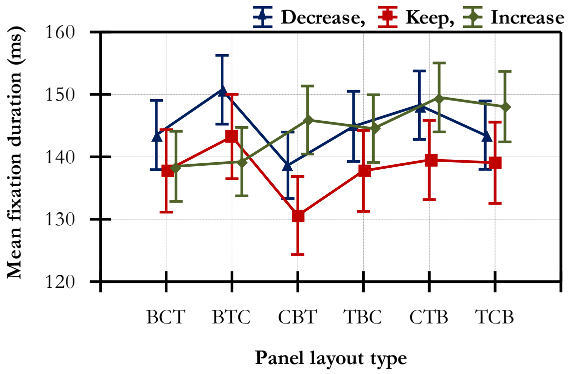

Fig. 8. *Panel layout type* × *Correct response type* interaction effect on mean fixation duration fixation duration for whole panels. $F_{Layout \times CRT}(10, 16\,786) = 1.9$, $p = 0.037$, $\eta^2 = 0.00115$.

b) Number of fixations

The influence of *Gender*, *Layout*, and *Correct response type* effects on a mean number of fixations was formally verified by a three-way Anova. Due to the limit regarding the number of custom stimuli in the BeGaze software and resulting from this exclusion of *Target velocity location*, the fixation count was aggregated by this factor. Thus, to make the interpretation of the results more intuitive, the cumulative number of fixations was divided by three to get the average fixation count per one panel version.

The analysis revealed a strong, statistically significant impact of the *Correct response type* factor ($F_{CRT}(2, 792) = 68$, $p < 0.0001$, $\eta^2 = 0.15$). The remaining two effects along with all interactions occurred to be irrelevant ($F_{Gender}(1, 792) = 2.1$, $p = 0.15$; $F_{Layout}(5, 792) = 1.03$, $p = 0.4$). The values of a mean number of fixations for the significant effect are graphically presented in Fig. 9. The data shows that the lowest mean number of fixations was observed for situations when subjects were to keep the current velocity unchanged (5.35 ± 0.146 *MSE*) than when the *Decrease* (7.5 ± 0.146 *MSE*) or *Increase* (7.4 ± 0.146 *MSE*) buttons had to be selected. LSD Fisher's analyses confirmed significant differences ($\alpha = 0.0001$) between *Keep* vs. *Decrease* and *Keep* vs. *Increase* levels, and no difference between *Decrease* vs. *Increase* levels ($p = 0.64$).

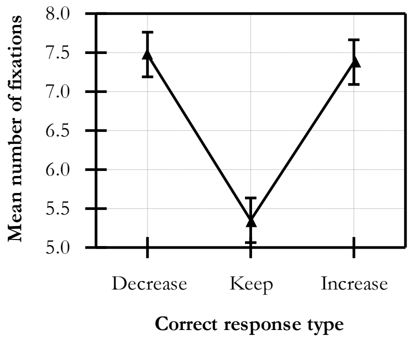

Fig. 9. *Correct response type* effect on mean number of fixations. $F_{CRT}(2, 792) = 68$, $p < 0.0001$, $\eta^2 = 0.15$.

c) Pupil diameters

Similar approach was applied to test the *Gender*, *Layout*, and *Correct response type* impact on the pupil diameter. This time, the three-way Anova showed meaningful effects of *Gender* ($F_{Gender}(1, 16\,679) = 519$, $p < 0.0001$, $\eta^2 = 0.03$) and *Correct response type* ($F_{CRT}(2, 16\,679) = 8.2$, $p = 0.0003$, $\eta^2 = 0.001$). The *Layout* factor ($F_{Layout}(5, 16\,679) = 1.3$, $p = 0.27$) and all interactions were insignificant ($\alpha = 0.1$). Significant factors are presented in Fig. 10 and 11. Fig. 10 demonstrates that in males the average pupil diameter was decidedly smaller (3.54 mm ± 0.0051 *MSE*) than in women (3.71 mm ± 0.0053 *MSE*).

As far as *Correct response type* is concerned, one may observe in Fig. 11 that the subjects for the *Keep* level exhibited the smallest average pupil diameter (3.6 mm ± 0.0071 *MSE*) while the highest values of the mean pupil size were recorded when the experimental task required to *Increase* the current velocity (3.64 mm ± 0.006 *MSE*). For the *Decrease* condition the value was in between *Keep* and *Increase* (3.62 mm ± 0.006 *MSE*). LSD Fisher's comparisons showed significant ($\alpha = 0.05$) differences between all levels of this factor.

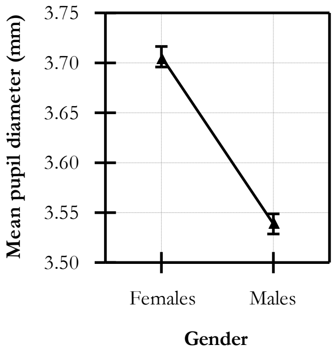

Fig. 10. *Gender* effect on mean pupil diameter. $F_{Gender}(1, 16\,679) = 519$, $p < 0.0001$, $\eta^2 = 0.03$.

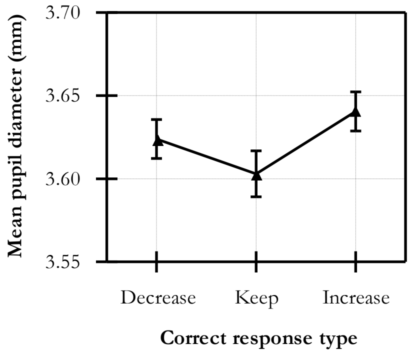

Fig. 11. *Correct response type* effect on mean pupil diameter. $F_{CRT}(2, 16\,679) = 8.2$, $p = 0.0003$, $\eta^2 = 0.001$.

2) Analyses of variance for panel components

In order to analyze what was the subjects' visual behavior in relation to the main panels' components, a series of two-way Anovas was applied. Together with a *Panel component type*, a *Gender* effect was included to test if an interaction exists.

a) Fixation durations

The examination of the influence of *Panel component type* and *Gender* on the fixation duration by Anova showed significance of both examined factors: $F_{Gender}(1, 16\,669) = 52$, $p < 0.0001$, $\eta^2 = 0.0031$; $F_{Components}(2, 16\,669) = 63$, $p < 0.0001$, $\eta^2 = 0.0076$. The interaction between these effects was not statistically meaningful ($F_{Gender \times Components}(2, 16\,669) = 1.03$, $p = 0.36$). The results regarding *Gender* are naturally identical to the Anova findings from the whole panel analysis. As to the panel *Components* factor, the Anova showed noticeably longest fixation durations for the *Control buttons* section (153 ms ± 1.41 *MSE*). The shortest mean was obtained for the *Current velocity* component (135 ms ± 1.1 *MSE*). The fixation duration mean for the *Target velocity* was in between (142 ms ± 1.2 *MSE*).

These data are graphically presented in Fig. 12. The outcomes of LSD Fisher' tests proved the statistical significance of differences between all factor conditions (α = 0.0001).

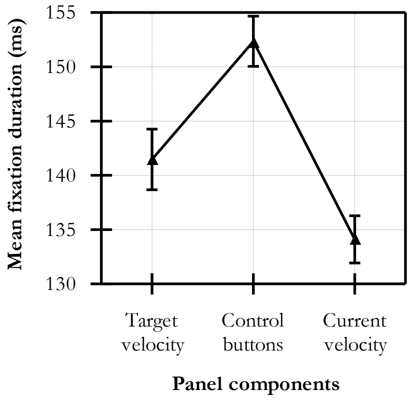

Fig. 12. *Panel component type* effect on mean fixation duration. $F_{Components}(2, 16\,669) = 63$, $p < 0.0001$, $\eta^2 = 0.0076$.

b) Number of fixations

The data for this analysis were prepared analogically to the fixation count analysis for the whole control panels. The findings obtained by (*Panel component type*) × (*Gender*) Anova confirmed significant influence of the *Panel components* effect on the mean number of fixations ($F_{Components}(2, 2425) = 150$, $p < 0.0001$, $\eta^2 = 0.11$).

Average values of the number of fixations depending on the *Panel components* factor are demonstrated in Fig. 12. The lowest mean fixation count was recorded for the *Target velocity* case (1.7 ± 0.043 *MSE*) whereas the biggest number was registered for *Current velocity* (2.7 ± 0.042 *MSE*). The average fixation count for the *Control buttons* section amounted to 2.5 ± 0.043 *MSE*. Post-hoc LSD Fisher's tests showed statistically important differences between all pairs of levels at $\alpha = 0.0005$.

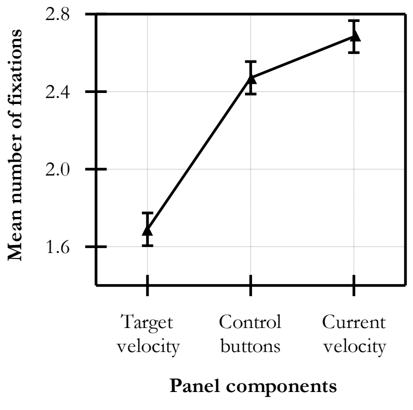

Fig. 13. *Panel component type* effect on mean number of fixations. $F_{Components}(2, 132) = 25$, $p < 0.0001$, $\eta^2 = 0.27$.

Gender along with the *Gender* × *Components* interaction occurred to be statistically irrelevant: $F_{Gender}(1, 2425) = 2.5$, $p = 0.11$; $F_{Gender\times Components}(2, 2425) = 1.9$, $p = 0.15$.

c) Pupil diameters

Another two-way Anova examined the impact of *Gender* and *Panel component type* on the pupil diameter. The obtained data provide evidence for the meaningful influence of the *Panel component type* factor on pupil diameter means ($F_{Components}(2, 16\,669) = 70$, $p < 0.0001$, $\eta^2 = 0.008$). The data presented in Fig. 14 show that on average the smallest pupil diameter was for the *Current velocity* level (3.59 ± 0.0057 *MSE*), slightly bigger for *Target velocity* (3.60 ± 0.0074 *MSE*), and markedly higher for the panel section *with Control buttons* (3.68 ± 0.0060 *MSE*). LSD Fisher's comparisons revealed important between *Current velocity* vs. *Control buttons* and *Target velocity* vs. *Control buttons.* The discrepancy between *Current* and *Target velocity* was insignificant ($\alpha = 0.05$). The outcome regarding *Gender* is the same as in the Anova results from the whole panel analysis. No interaction between *Panel component types* and *Gender* was observed: $F_{Gender\times Components}(2, 16\,669) = 2.2$, $p = 0.12$.

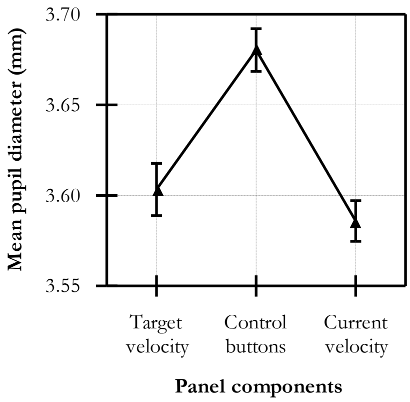

Fig. 14. *Panel component type* effect on mean pupil diameter. $F_{Components}(2, 16\,669) = 70$, $p < 0.0001$, $\eta^2 = 0.008$.

IV. Discussion and Conclusion

The presented study examined the impact of mutual location of panel basic components (control buttons, current and target velocities), target velocity location on the clock-face speedometer, and a correct response type on usability. The usability analysis included both typical performance measures along with spatiotemporal oculographic data.

The obtained data showed that women executed the experimental tasks slower than men and committed comparable number of mistakes. This outcome is partly consistent with tasks performed in some previous works (e.g. [4], [21], where women were slower than men but were more accurate. The overall number of errors in this study was more than twice as big as in [21] which indicates a more difficult task. Thus, it seems that the increase in cognitive workload necessary to complete a task eliminated better results of women effectiveness.

The lack of panel layout influence on task completion times was somewhat surprising since one could predict longer times for arrangements where the target and current sections were separated by the control buttons section. However, this was not the case which suggests that the task mental workload seemingly stronger influenced the total trial execution time than time needed for moving the eye balls between these panel components. Furthermore, in the experimental task necessary mouse movement were relatively short, and probably were performed during making the decision and searching for the correct button, and thus, did not affected the efficiency results.

Tasks completion times depended strongly on the *Correct response type* factor. For cases when the *Decrease* or *Increase* response was expected the task lasted on average about 30% longer than for the *Keep* option. Interesting is also the fact of committing drastically more errors for the *Increase* and *Decrease* in comparison with *Keep* where the number of errors was symbolic (merely 0.33%). Regarding effectiveness, the thought-provoking is the finding showing markedly bigger proportion (by more than 4%) of incorrect responses for the *Decrease* condition versus *Increase*, all the more that the completion times for these two levels were similar.

The oculographic data analysis presented in the current study was confined to the simplest and the most common measures. They included fixation count, fixation durations, and pupil diameters. The analysis of overall dwell times for consecutive experimental conditions was not conducted since they corresponded fully to the total task completion times discussed earlier.

The conducted analyses provide evidence of significant impact of three examined factors differentiating stimuli on average fixation times. As early as in 1935, Buswell [6] noticed that longer fixations are associated with more difficult the visual task is. Similar findings were also provided in later studies (e.g. [13] or [17]). In view of these findings, longer mean fixation durations for females than males show that the experimental task was more difficult and required more attention for women than men. Similar confirmation of the efficiency data analysis by visuospatial results is observed for the *Correct response type* effect where the mean values of task completion times exhibit the same pattern as average fixation durations.

On the other hand, the significant differentiation of mean fixation times by (*Panel layout*) and (*Panel layout* × *Target velocity location*) factors was not observed for the task completion times. Thus, visual activity data in this respect provided additional information on how the overt attention was distributed during task executions. It seems that, in general, arrangements with target velocity section located at the bottom of the panel required the least attentional effort whereas for options with control buttons at the panel bottom the attentional cost was the biggest. Additional interesting observations can be drawn from the *Panel layout type* × *Correct response type* interaction (Fig. 8) where the main source of the interaction effect was caused by the *Increase* option results. These qualitatively different outcomes might have resulted from a stronger emotional attitude indicated by bigger pupil dilations [5] registered for *Increase* stimuli than for other conditions (Fig. 11). This effect may stem from people's real life experiences where increasing the vehicle velocity can induce arousal. Larger pupil sizes were also observed for women vs. men. Such a result suggests either different emotional response [5] or bigger task difficulty for females than males [18].

A more detailed insight into the subjects' visual behavior was achieved by analyzing oculographic parameters for panel components across all conditions. Results regarding fixations revealed shorter but more numerous fixations for the *Current velocity* section as opposed to decidedly fewer but longer mean fixation durations for the *Target velocity section*. This finding might result from different presentation of the target and current velocity, and suggests different visual characteristics for them. The information demonstrated by a digital speedometer required longer mental processing accompanying by smaller number of attention shifts which could be caused by a more compact data presentation. On the other hand, the analogue (clock-face) speedometer seemed to be easier processed, therefore average fixation times were shorter. However, since the area covered by this section was twice as big as in the digital case, the number of necessary fixations to identify displayed velocity was bigger. Interestingly, mean pupil sizes results showed strongest arousal for the Control buttons panel section. This may be associated with the excitement accompanying the moment of making a decision.

In comparison with data presented by [13] regarding good and poor panels' designs, overall, the current study experimental tasks required less attention (142ms vs. 401ms) but required markedly bigger number of fixations (6.75 vs. 2.35). In [13], the number of fixations was related to the number of objects that a participant needed to process before finding the appropriate one. In the present experiment the fixations count is bigger probably due to the repetitive shifting between panels components since the functional grouping of target objects did not exist.

Some of the obvious limitations regarding the presented study are concerned with the eye tracking methodology applied. Analyses and interpretations such data relate only to the so called overt attention [12] while the covert one is registered. Moreover, oculographic data refer only to foveal vision and the peripheral field of view in not taken into account. Due to the article length restrictions, not all possible visual activity measures were analyzed. Some additional parameters could include scan path lengths, relations between the number of saccades and fixations etc. Given the nature of the experimental task which required comparisons between two types of information, it seems that an analysis of transitions between AOIs or even their parts could also be especially interesting. In this regard, a further extension of the current study is possible by attempting to model the obtained eye tracking data by means of the Hidden Markov Process. Some additional comparative analyses with similar control panels in a real environment may also be conducted in future research.

Despite the abovementioned limitations, the current research provides valuable contribution by extending our theoretical knowledge not only about the influence of examined factors on control panel design but also about the nature of visual activity patterns accompanying simple monitoring tasks. The presented findings may also be used by ergonomic practitioners and software developers while designing digital versions of various types of control panels.

Acknowledgment

The work was partially financially supported by Polish National Science Centre Grant No. 2011/03/B/ST8/06238. The experiment was conducted by means of the eye tracking system made available by the Laboratory of Information Systems Quality of Use which is a part of a BIBLIOTECH project cofounded by the European Union through the European Regional Development Fund under the Operational Programme Innovative Economy 2007-2013.

I thank Jerzy Grobelny for his for his invaluable support and helpful and stimulating discussions about this work.

References

[1] Andreoni, G., Mazzola, M., Zambarbieri, D., Forzoni, L., D'Onofrio, S., Viotti, S., Santambrogio, G. C., Baselli, G. (2013). Motion analysis and eye tracking technologies applied to portable ultrasound systems user interfaces evaluation. In *2013 International Conference on Computer Medical Applications (ICCMA)* (pp. 1–6). doi:10.1109/ICCMA.2013.6506170.

[2] Barbieri, L., Angilica, A., Bruno, F., & Muzzupappa, M. (2013). Mixed prototyping with configurable physical archetype for usability evaluation of product interfaces. *Computers in Industry*, *64*(3), 310–323. doi:10.1016/j.compind.2012.11.010.

[3] BeGaze Manual, Version 3.6, February (2016), SMI SensoMotoric Instruments, document number: 091222-P-1400-001-000-A.

[4] Blatter, K., Graw, P., Munch, M., Knoblauch, V., Wirz-Justice, A., & Cajochen, C. (2006). Gender and age differences in psychomotor vigilance performance under differential sleep pressure conditions. Behavioural Brain Research, 168(2), 312–317. doi:10.1016/j.bbr.2005.11.018.

[5] Bradshaw, J. (1967). Pupil Size as a Measure of Arousal during Information Processing. *Nature*, *216*(5114), 515–516. doi:10.1038/216515a0.

[6] Buswell G.T., (1935). How People Look at Pictures. A Study of the Psychology of Perception in Art, The University of Chicago Press, Chicago, IL.

[7] Campos, J. C., & Harrison, M. D. (2008). Systematic Analysis of Control Panel Interfaces Using Formal Tools. W T. C. N. Graham & P. Palanque (Ed.), *Interactive Systems. Design, Specification, and Verification* (pp. 72–85). Springer Berlin Heidelberg. Doi: 10.1007/978-3-540-70569-7_6.

[8] Crowe, E. C., & Narayanan, N. H. (2000). Comparing Interfaces Based on What Users Watch and Do. In: *Proceedings of the 2000 Symposium on Eye Tracking Research & Applications* (pp. 29–36). New York, NY, USA: ACM. doi:10.1145/355017.355022.

[9] Duchowski, A. T. (2002). A breadth-first survey of eye-tracking applications. *Behavior Research Methods, Instruments, & Computers*, *34*(4), 455–470. doi:10.3758/BF03195475.

[10] Duric, Z., Gray, W. D., Heishman, R., Li, F., Rosenfeld, A., Schoelles, M. J., Schunn, C., Wechsler, H. (2002). Integrating perceptual and cognitive modeling for adaptive and intelligent human-computer interaction. *Proceedings of the IEEE*, *90*(7), 1272–1289. doi:10.1109/JPROC.2002.801449.

[11] Fidler, R., & Johnson, M. (2016). Human Factors Approach to Comparative Usability of Hospital Manual Defibrillators. Resuscitation, 101, 71–76. doi:10.1016/j.resuscitation.2016.01.029.

[12] Findlay, J.M. and Gilchrist, I.D. (2003), Active vision. The psychology of looking and seeing, New York: Oxford University Press.

[13] Goldberg, J. H., & Kotval, X. P. (1999). Computer interface evaluation using eye movements: methods and constructs. International Journal of Industrial Ergonomics, 24(6), 631–645. doi:10.1016/S0169-8141(98)00068-7.

[14] Grobelny J., Michalski R. (2015). Comparative analysis of regular grid based algorithms in the design of graphical control panels. In: M. Kurosu (Ed.): Human-Computer Interaction: Design and Evaluation, Part I, HCII 2015, LNCS 9169, pp. 332–339, Springer. doi: 10.1007/978-3-319-20901-2_30.

[15] ISO CD 9241-11: Ergonomics of human-system interaction – Part 11: Usability: Definitions and concepts (2015).

[16] Jacob, R. J. K., & Karn, K. S. (2003). Commentary on Section 4 - Eye Tracking in Human-Computer Interaction and Usability Research: Ready to Deliver the Promises. In: R. Radach & H. Deubel (Eds.), *The Mind's Eye* (pp. 573–605). Amsterdam: North-Holland.

[17] Just, M. A., & Carpenter, P. A. (1976). Eye fixations and cognitive processes. *Cognitive Psychology*, *8*(4), 441–480. doi:10.1016/0010-0285(76)90015-3.

[18] Kahneman, D., & Beatty, J. (1966). Pupil Diameter and Load on Memory. *Science*, *154*(3756), 1583–1585. doi:10.1126/science.154.3756.1583.

[19] Kwahk, J., & Han, S. H. (2002). A methodology for evaluating the usability of audiovisual consumer electronic products. Applied Ergonomics, 33(5), 419–431. doi:10.1016/S0003-6870(02)00034-0.

[20] Liu, K.-H., Hwang, S.-L., Hsieh, M.-H., Max Liang, S.-F., & Chuang, C.-F. (2016). Systematic layout planning in human–system interface: An evaluation of alarm displays with spatial proximity for accidents diagnosis of advanced boiling water reactor. *International Journal of Industrial Ergonomics*, *51*, 30–42. doi:10.1016/j.ergon.2014.12.014.

[21] Michalski R. (2008), The effects of panel location, target size, and gender on efficiency in simple direct manipulation tasks. [In:] I. Pavlidis (red.), Advances in Human-Computer Interaction, Austria, Vienna: In-Tech Education and Publishing, pp. 11–22.

[22] Pala, S., Schnell, T., Becklinger, N. L., Giannotti, C., Sun, B., Tanaka, H., & Shimonomoto, I. (2011). *Adaptation of the Cognitive Avionic Tool Set (CATS) into Automotive Human Machine Interface Design Process*. SAE Technical Paper no. 2011-01-0594. doi:10.4271/2011-01-0594.

[23] Renshaw, J. A., Finlay, J., & Webb, N. (2006). Getting a Measure of Satisfaction from Eyetracking in Practice. In: *CHI '06 Extended Abstracts on Human Factors in Computing Systems* (pp. 1723–1726). New York, NY, USA: ACM. doi:10.1145/1125451.1125772.

[24] Sargent, T. A., Kay, M. G., & Sargent, R. G. (1997). A Methodology for Optimally Designing Console Panels for Use by a Single Operator. *Human Factors: The Journal of the Human Factors and Ergonomics Society*, *39*(3), 389–409. doi:10.1518/001872097778827052.

[25] Smeets, J. B. J., & Hooge, I. T. C. (2003). Nature of Variability in Saccades. *Journal of Neurophysiology*, *90*(1), 12–20. doi:10.1152/jn.01075.2002.

[26] Tan, H., Boon, M. Y., & Dain, S. J. (2014). The visibility of controls and labels on electronic devices and their suitability for people with impaired vision. *Work (Reading, Mass.)*, *47*(3), 309–317. doi:10.3233/WOR-131769.