HAL open science

Pupil-Assisted Target Selection (PATS)

Christoph Strauch, Jan Ehlers, Anke Huckauf

To cite this version:

Christoph Strauch, Jan Ehlers, Anke Huckauf. Pupil-Assisted Target Selection (PATS). 16th IFIP Conference on Human-Computer Interaction (INTERACT), Sep 2017, Bombay, India. pp.297-312, 10.1007/978-3-319-67687-6_20 . hal-01717224

HAL Id: hal-01717224

https://inria.hal.science/hal-01717224v1

Submitted on 26 Feb 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License

Pupil-Assisted Target Selection (PATS)

Christoph Strauch, Jan Ehlers, and Anke Huckauf

Department of General Psychology, Ulm University, Ulm, Germany
{christoph.strauch, jan.ehlers, anke.huckauf}@uni-ulm.de

Abstract. Physiological signals such as pupil size changes promise improvements for human-computer-interaction. However, the pupil response is known to be rather slow and unspecific. This hinders its application in target selection up to now. Nevertheless, there are indications for fast diameter changes accompanying cognitive processing already at early stages so that pupil effects can even precede psycho-motor activity. Building on these findings, we investigated the potential of short-latency pupil size changes for improving target selection in a search and select task. Pupil assisted target selection (PATS) was shown to be competitive to both, purely pupil-based and to dwell-time based selection modes in regard to selection times, but at the cost of more false positives than for a dwell-time approach in a search and select task. This demonstrates the usefulness of PATS as a means for target selection. The observed pupil dynamics correspond to early signal courses in basic research. Pupil dynamics also suggest room for further improvements of the integrated concept of pupil-assisted target selection.

Keywords. Gaze-based interaction; cognitive pupillometry; physiological computing; eye-tracking

1 Introduction

1.1 Physiological Interaction

Current multimodal systems vary in respect of number and kind of input channels. The growing field of physiological computing extends parameterization by providing human-computer interfaces (HCI) with physiological information [1]. Physiological data includes information on the psychophysiological state of the user and is usually applied to derive emotional features and to make systems aware of a user's (cognitive) condition. However, one vision of physiological computing frameworks is that the corresponding information (e.g., pupil dynamics, electrodermal activity changes or heart rate variability) can even be transferred into active control signals. Explicit commands (voice commands, gestures or keystrokes) could be supported through a broader range of signals or would become obsolete to a certain degree [2]. Furthermore, directly accessing input channels via implicit information may extend communication to a variety of devices for which standard interfaces are not suitable. For example, virtual-reality headsets could be operated by tracking pupil size. Including physiological data might

also lead to more inclusive interfaces which can be used by all, healthy persons as well as severely disabled people [3].

Recent studies on physiological computing include a variety of physiological variables that contain relevant information on a user's cognitive and affective state [3]. However, the depicted signals are usually diffuse and emerge as unspecific functions of autonomic activity changes. E.g. pupil size changes may indicate the amount of cognitive load just as various levels of affective processing [4]. This low specificity has so far prevented a straight forward application of physiological variables in HCI. Furthermore, most parameters are characterized by rather long latencies that prevent from a direct application. However, changes in pupil diameter constitute one of the fastest psychophysiological variables, with latencies below 500 ms after stimulus onset [5]. There is also evidence that diameter changes not only accompany, but to some extent precede psychomotor processing [6, 7]. Early pupil size changes that reliably occur during target selection may therefore be used as a predictor and thus as an implicit command for target selection. The current work introduces a target selection procedure that combines gaze information and pupil dynamics to implement a new concept of eyes-only interaction in HCI.

2 Related Work

Pupil size is determined by two antagonistic muscle groups, governed by the sympathetic and parasympathetic division of the autonomous nervous system. An increase in physiological activity is associated with an enlargement of pupil size, whereas low autonomic arousal correlates with reduced pupil diameter. Thus, pupil size is not exclusively adjusted by changes in illumination but also provides a valid measure of a subject's cognitive and affective state [8]. The amount of mental workload is reflected in an enlargement of pupil diameter [9] such as the dilation in response to emotionally arousing sounds [5]. Attentional processes may also be investigated by deconvoluting the pupillary signal [10] just as decision processes in binary tasks [7, 11] or the decision to click on a website [12]. Even hunger is reported to affect baseline pupil sizes [13]. Furthermore, detecting and recognizing perceptual stimuli is known to evoke pupil dilations [8]: The so-called "navigational intent", refers to a state in which subjects try to get an overall picture of an image. During this period, pupil size is lower compared to phases of "informational intent", in which the user searches for a particular object [14]. To the present day, pupil size changes are mainly used as a passive indicator for cognitive processing; however, there is evidence that the corresponding dynamics can be deliberately influenced and may serve as an active information channel [4].

Pupil sizes are obtained with camera systems; the associated dynamics are reported to entail latencies of about 0.5 s. As reaction to emotionally charged sounds, a maximum pupil dilation about 2-3 s after stimulus onset is reported [5]. The human pupil size varies from under 1 mm to over 9 mm, psychologically evoked diameter changes are rather small and reported to amount up to 0.5 mm [8]; however, [4] report pupil size changes of up to 1 mm in an active biofeedback application.

Changes in pupil diameter are therefore of central interest for physiological computing systems. However, underlying mechanisms are not always clearly distinguishable and may reflect similar operating principles. Hence, it is necessary to disentangle the associated dynamics and to identify specific signal courses that can reliably be assigned to definable cognitive processes.

2.1 Pupil Size Changes as an Active Information Channel in HCI

In HCI, pupil sizes changes are usually applied to derive mental effort on-line [15]; only few studies are known to utilize pupil dynamics as an active input channel for HCI [16–18].

A recently introduced pupil-based interaction concept [17] builds on the finding that pupil diameter can be subject to voluntary control albeit via cognitive strategies and could be used as an input mechanism to HCI [4, 19, 20]. Here, the induction of cognitive load enabled locked-in patients to communicate on the basis of changes in cognitive load. Different arithmetic problems were successively displayed on a screen in conjunction with the words “yes” or “no”. Participants select the response options by processing the displayed task and hereby increase pupil diameter via cognitive load in one of the two phases. The slope of a regression then indicated which answer was envisaged to be selected. The reactions to questions with obvious solutions enable to check for error rates. This approach is especially promising, since it may be applied to clinical subjects outside the laboratory. However, error rates are still high and, more importantly, selections are only binary and take several seconds to be completed. Moreover, data could only be analyzed post-hoc so that live communication was not possible and explicit mental arithmetic is necessary to operate the system [17]. [16] make use of pupil size changes during target selection. In this setting, targets slowly oscillate with different frequencies, whereas pupil dynamics adapt to the flickering source. Here, the number of different options is limited by the number of frequencies in the lower range that can reliably be distinguished from each other. Similar to the aforementioned, this selection mechanism takes several seconds to complete a selection.

Taken together, current approaches that apply pupil dilation for active target selection are associated with overly long selection times and partially high error rates. Hence, pupil-based target selection seems to be unsuitable for real time computer input in end-user applications. However, for scenarios, such as the operation of cognitively demanding tasks, also active pupil size manipulation should be considered [18].

2.2 Selection-Associated Pupil Size Changes

However, besides the mentioned slow changes in pupil size with mental load, also motor actions and their precedent decision processes are reported to be accompanied by pupil size changes [6, 7, 21]. According to [7], [21], and [12], one might expect that the decision to select a certain object affects pupil size: more concretely, when a user fixates an attracting target, pupil size should be expected to even precede or occur within few hundreds of milliseconds. In line with these findings, finding a target during visual search is associated to a fast dilation in contrast to fixations off-target [22, 23]. However, in a machine learning approach, pupil size variation only slightly contributes to

post-hoc classification of user intent in addition to other variables derived by eye-tracking [24]. For objects, which are not considered as potential target by the user, pupil size should be expected to vary within its standard noise. Up to now there is no research examining such early pupil size in applied scenarios. Still, such early signal changes may provide a new approach for pupil-based target selection. Based on these signal dynamics, we developed an eyes-only interface for a pupil assisted target selection (PATS) mechanism that aims to equally reduce selection times and error rates.

2.3 Research Aims

- Investigate the potential of using pupil size changes to improve target selection at an early stage
- Compare PATS with standard gaze-based dwell-time approaches
- Assess early pupil dynamics during target selection

3 Methods

3.1 Pupil Assisted Target Selection (PATS)

PATS integrates pupil size changes with information on gaze behavior as two central input variables. Implemented as a usual dwell-time approach, target selection may either be achieved by a simple fixation duration or else through pupil enlargement beyond a predetermined threshold, whichever is reached first. The selection process was thereby accompanied by feedback. Fixated objects were consistently highlighted. However, PATS was assessed in two versions, differing in the pupil feedback provided for the user. As a comparison for PATS, a purely dwell-time selection approach was used in which users had to fixate a target for a eye-typing novice dwell-time (1000 ms) [25]. In addition, a pupil size-based selection mechanism was implemented, in which a calibrated increase in pupil size over 667 ms was the criterion for object selection. Window length was chosen based on the dynamics reported in [7] and [22]. In both papers, considerable increases are registered within lower than 1 s. This length corresponded to 40/60 Frames. The deviation between beginning and end of the 667ms was calculated continuously and started when a new trial was presented. That is, objects did not have to be fixated. Hereby, also fixation slightly preceding dilations [7, 22] could be used. For all modes, an initial dwell-time of 500 ms was necessary to enable selection.

3.2 Task

The system was evaluated employing a search and select task: several targets are arranged in a circle. One of the circularly arranged objects serves as target. The target is defined via its similar shape to a centrally placed object. Figure 1 illustrates the task layout. In the depicted example, the object at 8 o´clock corresponds to the central object and would thus have to be selected.

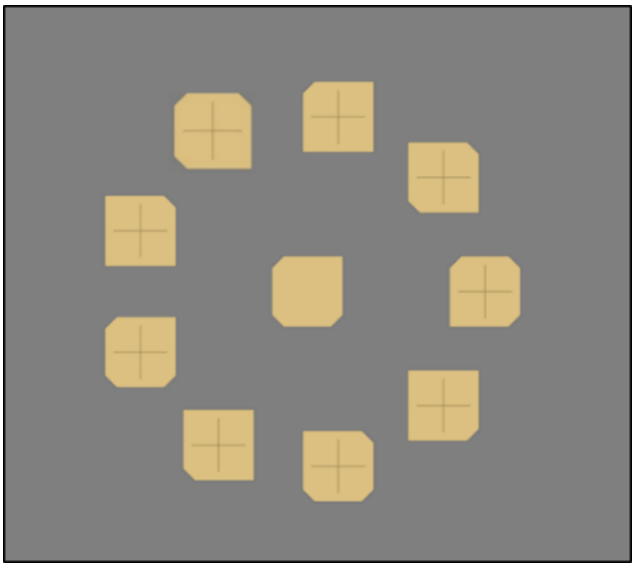


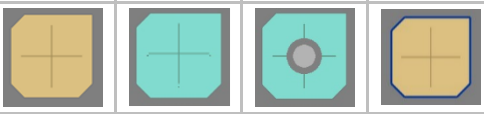
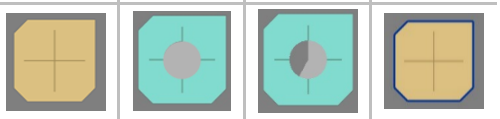
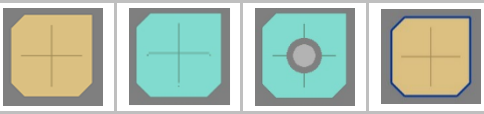
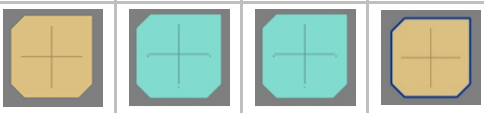
Fig. 1. Search and select task with nine circularly arranged objects including one target corresponding to the reference object in the center.

Error rates and selection times were assessed in order to determine effectiveness and efficiency, respectively. Error rates are computed as the sum of false positive non-intended selections and time-out errors. A trial was counted as time-out when users did not select a target within 20 s. The selection time is the time from first fixating the intended target until selection. User satisfaction was assessed with the "system-usability scale" (SUS), ranging from 0 (no user satisfaction) to 100 (perfect user satisfaction). The SUS also measures learnability by determining how difficult it is to adapt to a system's set of rules [26].

3.3 Feedback and Selection Modes

Selection performance using PATS was assessed in two ways that varied with regard to the kind of feedback. Both approaches were compared to a purely pupil-based and a purely fixation (dwell-time)-based selection mode. All four concepts are illustrated in Table 1.

Table 1. Presentation of target and feedback for the four selection modes. Feedback circles were dynamically filled until the target was selected.

Objects	Target not fixated	Target fixated	Target fixated >500 ms	Target selected
Pupil Size				
Dwell-Time				
PATS				
PATS Reduced FB				

When fixated, every object was highlighted in light blue. Unattended objects were displayed in yellow as depicted in Figure 1. For the selection mode purely based on pupil size changes, a dark grey circle appeared 500 ms after a target was fixated. A bright grey circle visualizing the criterion for selection was shown within the fixated target. The size of this circle was meant to illustrate the increase in pupil size over 667 ms required for selection, a threshold which was determined during calibration before each block. Within this circle, a light grey circle dynamically represented the current increase in pupil size pupil diameter within this time window. As soon as the outer circle was filled, the target was selected. Subsequently, the associated target was framed with a blue line to indicate a successful selection for 2 s. Each trial was followed by a pause of 1 s.

The sequence of events in the dwell-time selection mode was the same as for the purely pupil-based selection, except the following differences: As soon as gaze entered the object, a light grey circle was displayed in the object. It was filled clockwise by a brighter circle. After fixating the object for a period of 1000 ms, the light circle was completely filled and the object was selected.

The PATS selection mechanism combines the purely pupil-based and the dwell-time based approach. A static light grey circle appeared within the dark grey circle. For PATS with reduced feedback, only the gaze-based highlighting was applied. The sequence of events is illustrated in Table 1 separately for all four selection modes.

3.4 Apparatus

Pupil size and gaze were tracked using a SMI iView XTM Hi Speed 1250 eye-tracker (SensoMotoric Instruments GmbH, Teltow) positioned 65 cm from the screen (BenQ

XL2720Z 1920*1080 Pixels, 60 Hz). For the experiment, the eye tracker was also set to a sampling rate of 60 Hz. Data were collected using iViewX 2.8.43 software. Following the manufacturer, tracker precision is better than 0.01° visual angle, precision for pupil size estimation is 0.01 mm or better. Diameter was calculated using a pixel to mm conversion by the manufacturer. A chin rest secured the constant distance and position of the seated participants' heads and eyes to both, eye-tracker and screen. The task, presentation of visual stimuli, the underlying selection modes, as well as data savings, were implemented using PsychoPy version 1.81.02 [27].

3.5 Sample

24 users participated in the experiment, all of them were students at Ulm University (female = 18, M_{Age} = 24.08). All users reported normal or corrected to normal vision. All participants reported to not having consumed drugs or medication prior to the experiment, no neuronal diseases or traumatic brain injuries were reported; and all participants were asked to stay absent from coffee for at least one hour prior to the experiment. Participants signed an informed consent and took part in the study on a voluntary basis. They received either course credits or a present containing a set of glittering unicorn-stickers and wiggle eyes as a reward.

3.6 Procedure

Experiments were carried out in a laboratory with constant brightness of 38.0 lx obtained at the eyes' position. After having signed the informed consent, users filled a questionnaire to assess demographic variables and a confidential questionnaire assessing potential traumatic brain injuries, the usage of illegal substances and neuronal diseases. This took about ten minutes. The users were then instructed about the search and select task using a printed screenshot. Users were informed that looking at the objects they intend to select was a prerequisite for selection. However, it was not explained how exactly a selection could be performed; meaning that the role of pupil size changes was not mentioned. In addition, participants were told that they will use different selection modes and were explained that they would have to evaluate them after the tasks. Additional questions were answered if occurred.

Users took a seat in front of the monitor in the eye-trackers´ chin rest. It was ensured that pupils and the corneal reflexes could be registered reliably in all gaze positions. Then the automatic SMI calibration started: 13 points were displayed systematically on the screen to calibrate the gaze position-mapping. Another 13 points appeared in order to validate the gaze-mapping. Users could access a keyboard with their right hand, which enabled to push ← and →. After calibration of the eye tracker, threshold for pupil-based selection was calibrated. Therefore, participants had to complete thirty trials of the search and select task employing the pupil selection mechanism. It was constantly determined whether pupil diameter enlarged for more than 0.2 mm within the foregoing 667 ms. If so, targets were selected; however, only after a minimum dwell-time of 500 ms. Whenever a target was selected correctly, the criterion diminished by 2 %. In case of a false positive selection, the criterion increased by 5 %, in case of a

time-out error, the criterion diminished by 5 %. The resulting criterion after thirty trials was then further used during the pupil-based or assisted target selection.

After calibrations, users performed twenty trials of the search and select task with each of the four selection modes. The serial order of the modes was fully permutated in six sequences; each sequence was completed by four participants. The PATS mode with reduced feedback was consistently applied last, since it was assumed that subjects needed training in order to be able to work with an input mode which displays only little feedback. After completing a selection mode, a short information was given that the next run is starting soon so that participants were able to distinguish the different concepts from another. Within each selection mode twenty trials of the search and select task with randomized target shapes were to be performed. After each target selection, a screen was presented: "Wolltest Du dieses Objekt auswaehlen? (nein = ←, ja = →). (Did you intend to select this object (no = ←, yes = →))". If 20 seconds were exceeded during a selection process, the respective trial was aborted, and "time out error" was displayed on the screen. Such a trial was counted as time-out. After pushing either ← or →, the next trial started after a one second-break. After three blocks with varying order of selection modes, the reduced-feedback selection mode was performed as a last condition. Finally, participants were asked to answer one SUS for each selection method and received their reward.

4 Results

Twentyfour users completed the study. Since the search and select task comprised twenty trials for each selection mode, error rates and selection times were calculated on the basis of 480 trials per selection mode. For PATS, 45.63 % of selections were performed using the pupil size selection mechanism and 54.37 % using dwell-time. For PATS with reduced feedback, 40.83 % of targets were selected with the pupil, while 59.17 % were selected by dwell. A summary of key usability parameters is depicted in Table 2. All dependant variables were analyzed employing ANOVAs for repeated measures. Whenever sphericity could not be assumed, Greenhouse-Geisser correction was applied. Post hoc comparisons were analyzed via Bonferroni corrected contrasts.

4.1 Error Rates

Among the four selection mechanisms, error rates were lowest for the dwell-time selection mechanism with .84 % errors (SE = 0.39 %). But, also pupil-based selection (5.41 %, SE = 1.34 %), PATS (6.88 %; SE = 1.27 %), and PATS with reduced feedback (5.41 %, SE = 1.47 %) produced error rates that still allow task operation. A significant difference between selection modes could be found ($F_{(2.21,\ 50.78)} = 6.24$; $p < .01$ $\eta^2_{part} = .21$). Post-hoc comparisons reveal that only dwell-time differed significantly from all other modes.

4.2 Selection Times

Selection times were longest for the pupil selection mechanism (1.48 s, *SE* = 0.15 s). However, it is important to note that observed average selection times for twelve of the 24 users are comparable to the dwell-time of 1000 ms or lower (Median = 1.13 s). PATS (0.82 s, *SE* = 0.01 s) and PATS with reduced feedback (0.84 s, *SE* = 0.01 s) produced lower selection times than the dwell-time of 1000 ms and merely pupil-based selection. Average selection times were comparable between subjects except for the pupil selection mechanism, for which individual differences were monitored. The distribution of selection times is depicted in Figure 2.

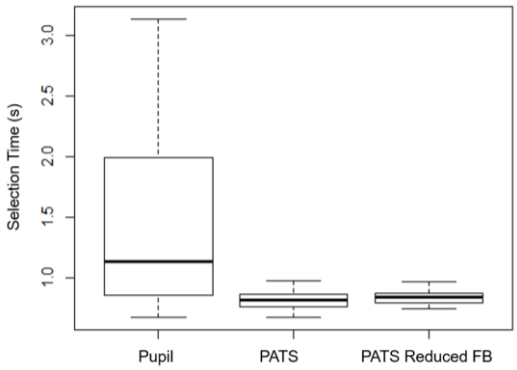


Fig. 2. Distribution of selection times for pupil based, PATS and PATS with reduced feedback selection modes.

Taken together, effectivity for purely pupil-based selection mode was lower than for dwell time selection. However, effectivity as well as efficiency were shown to be best for PATS and PATS with reduced feedback. All selection modes but PATS and PATS with reduced feedback differ from another significantly with regard to selection times (($F_{(1.02,\ 23.46)} = 17.46$, $p < 0.001$, $\eta^2_{part} = .43$).

4.3 User Satisfaction and Learnability

For usability, no significant differences in ratings were found ($F_{(2.25,51.66)} = 2.310$, $p = .10$, $\eta^2_{part} = .09$). However, descriptively, usability was rated worst when targets were selected utilizing the purely pupil selection mode (65.63, SE = 4.49) and almost similar for dwelling (76.56, SE = 3.85), PATS (75.26, SE = 4.13) and PATS without feedback (75.26, SE = 4.01). Following [28], these evaluations are all above the critical value for the 2nd quartile; dwelling, PATS and PATS with reduced feedback were even rated within the third quartile, indicating a good usability.

Similarly, learnability ratings did not differ significantly ($F_{(1.90,43.62)} = 2.77$, $p = .08$, $\eta^2_{part} = .11$). Descriptive statistics show that learnability was rated worst for PATS (72.92, SE = 5.30) and for the purely pupil-based selection mode (75.52, SE = 4.96). PATS with reduced feedback (79.69, SE = 4.87) and dwell-time based selection (83.54, SE = 3.65) received higher ratings.

Overall, usability (71.29, SE = 2.73) and learnability (77.99, SE = 4.02) scores reveal that users generally evaluated the selection modes as usable and learnable [28]. Effectivity, efficiency, and usability all indicate that PATS, especially PATS with reduced feedback, can be used for active target selection, although the implemented dwell-time based selection could not be outperformed.

Table 2. Key usability statistics separately for the four selection modes pupil, dwell, PATS and PATS with reduced feedback. Usability and learnability were assessed using the SUS; that is, the range of usability and learnability is 0-100, hereby 100 marks the best possible evaluation.

Selection Mode	False Positives *M*	Time-Out *M*	Selection Times *M*	Usability Rating *M*	Learnability Rating *M*
Pupil	4.58 %	0.83 %	1.48 s	65.63	75.52
Dwell	0.63 %	0.21 %	1.00 s	76.56	83.54
PATS	6.88 %	0.00 %	0.82 s	75.26	72.92
PATS Reduced FB	5.40 %	0.00 %	0.84 s	75.26	79.69

4.4 Signal Dynamics

In order to examine the dynamics of pupil signals prior to and subsequent to target selection, respective signals were further analyzed. Firstly, pupil signals in all trials in which selections were confirmed as unintended by the user were excluded in further analysis. Then, all trials in which targets were selected faster than 1.5 s after task onset were excluded. This was done, so that a proper in-task baseline could be registered and the second before target selection could be analyzed for all remaining trials. Data were normalized with a local baseline: the mean of the first 0.5 s of each trial was subtracted from each data point. For the purely pupil-based mode, 411 out of 480 trials met the criteria, 444 for dwelling, 406 for PATS, and 410 for PATS with reduced feedback. The signals were averaged separately by selection mode for each user. Users´ means are depicted in Figure 3. The graph gives mean pupil dynamics separately for selection modes, starting 1 s before entering the finally selected target until 2.5 s after entering the target. That is, the event of entering the finally selected target by gaze is set to 0 s.

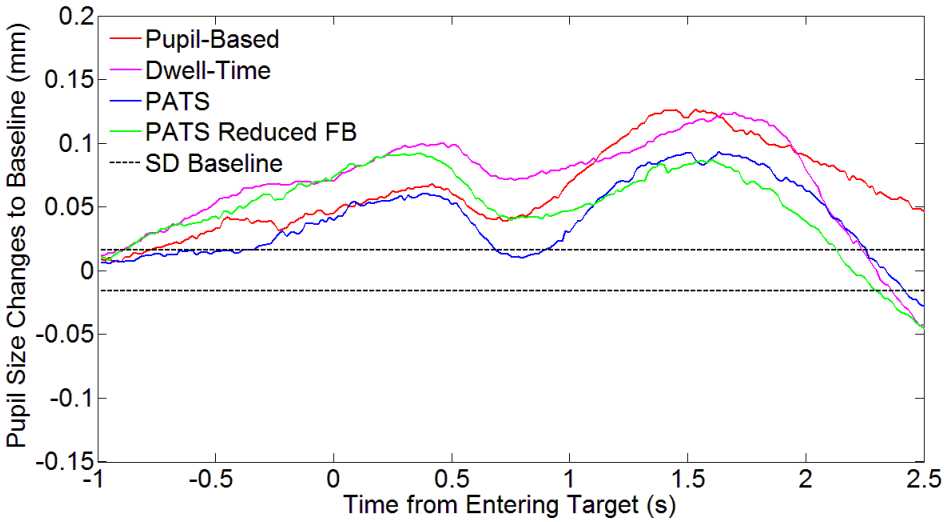


Fig. 3. Mean pupil size changes (differences to baseline) over time separately for selection modes. 0 s represents the time when the finally selected target was entered by gaze.

As can be seen in Figure 3, there seems to be a common course in the pupil dynamics for all selection modes: two local maxima, one at about 0.4 s, one at about 1.5 s, and one local minimum at about 0.8 s are recognizable for all pupil dynamics. This is even true for the dwell-selection mode, which did not require pupil size changes.

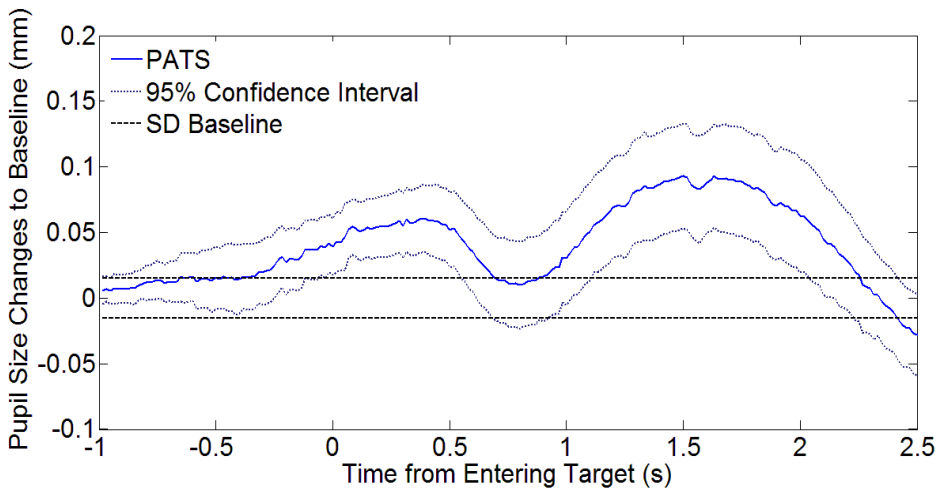


Fig. 4. Dynamics of mean pupil size changes (differences to baseline) during PATS. 0 s represents the time when the finally selected target was entered by gaze.

The maximum pupil response at about 1.5 s might be interpreted as a pupillary response to the load associated with the ongoing selection process, the feedback and the confirmation of successful selection. However, even before entering the target, mean pupil size rises considerably and independently from the selection mode. This signal course corresponds to the course reported in [7] but is even clearer. The signal shows a comparable dilation as in [22], however, additionally, a subsequent constriction is found. Figure 4 depicts the mean pupil size changes during the PATS mechanism together with a functional 95 % confidence interval calculated on the basis of 24 mean signal dynamics. Thus, for every averaged data point a confidence interval was calculated on the basis of the 24 mean-constituting data points. The functional confidence interval can thus also illustrate the significant changes to baseline occurring at about 0.4 s and 1.5 s [4].

Could the illustrated pupil size changes be elicited by the timing of the task, e.g. the foregoing pause and/or task onset alone instead of target selections? In order to investigate this question, we compared long and short trials, assuming that longer trials should be robust regarding carry over effects from foregoing events. For 980 out of 1651 trials, subjects needed 3 s or more until fixating the correct object (< 3 s: 671 trials). We find almost identical signal properties for long as for short task completion times entailing maxima at about 0.4 s and 1.5 s and a minimum at about 0.8 s.

5 Discussion

The current investigation aims to evaluate the potential of pupil size changes to actively support target selection. The integrative interaction concept PATS combines the benchmarking dwell-time approach with the newly developed concept of a purely pupil-based selection mode. Therefore, we assessed effectiveness, efficiency, and user satisfaction, as well as early pupil size courses that accompany object selection in a search and select task. Signal dynamics were analyzed from 1.5 s prior until 2.5 s after the first fixation of an intended target. The classical dwell-time selection mode was associated with highest accuracy rates; PATS (and PATS with reduced feedback) appeared to be less effective but entailed considerably shorter selection times. Key usability parameters were rated positively for all four evaluated selection modes. The tracked pupil sizes show early pupil size changes that accompany and precede the first fixations of intended selections.

In an interface that uses first settings for the crucial parameters, brightness, window length and deviation criterion, we assessed usability for PATS in a search and select task. Key-usability variables show that PATS is possible. Effectiveness, measured as percentage of incorrect answers, is best for the dwell-time approach; however, unsupported pupil-based selection, PATS, and PATS with reduced feedback also provide an acceptable usability, especially when taking the low experience with pupil-based selection approaches into consideration. With regard to efficiency, PATS and PATS with reduced feedback even outperformed target selection based solely on pupil size changes or dwelling. The median of average selection times for the purely pupil-based mechanism reveals that a comparable selection time could be achieved, although subjects who took much longer negatively distorted the mean. In addition, average error rates seem

to distinguish stronger between users than between trials within users. This was especially the case for the unsupported pupil-based approach and the PATS concepts: several users were able to select targets fast and reliably, whereas others had severe problems to consistently apply the required pupil size changes. But, speed and accuracy represent a trade-off. A comparison of error rates in dwell time and purely pupil-based selection mode to PATS and PATS with reduced feedback can only be carried out when average selection times roughly correspond to the dwell-time concept. Retrospectively, dwell-time should have been set slightly lower to the average selection time of PATS, given that almost no errors were observed. Error rates for dwelling might change if the dwell-time was reduced to the average selection times of PATS, this should be focused on in further studies in order to allow an adequate judgement of the competitiveness of dwell and pupil enhanced dwell-times.

Usability ratings did not differ significantly. This must not mean that they are in fact comparable. On a descriptive basis, usability of the dwell-time approach and the PATS concepts, were rated comparably good. However, merely pupil-based selection was evaluated worse than all other modes. Also in learnability, there were only minor differences, and dwell-time and PATS with reduced feedback receiving the highest ratings. One might speculate that the intuitive dwell-time feedback facilitates operability whereas feedback on pupil size changes appeared mostly too short and couldn't fulfill its actual function. User reports subsequent to the testing match the assumption that feedback in PATS and during unsupported pupil-based selection mode was disturbing and unhelpful. The rated learnability is on average higher than the usability evaluation. Subjects had to complete only 20 trials per selection mode, whereas one turn lasted for about 5 minutes. Given the short operation time, overall performance might be better in a longer experiment, since subjects were generally optimistic to further improve the ability to operate the system.

Pupil dynamics were comparable for each selection method. The obtained signal courses reveal two local maxima, 0.4 s and 1.5 s after target entry and irrespective of the applied selection mechanism. In PATS and purely pupil-based selection, an increase in pupil size can be expected as the selection principle also depends on the slope of the signal. However, the depicted signal courses can also be observed during dwell-time selection mode, for which pupil size changes are irrelevant for selection. Hence, one might suggest that this pupil dynamic accompanies target selection in general and can thus be used for further improvements and enlargements of PATS.

Concerning the second peak at about 1.5 s, one might assume that either the processing of the feedback information, and/or of the successful selection, might have caused this pupil reaction. But, there is also a first peak at about 0.4 s revealing a considerable increase in pupil size even before the target is fixated. In order to use respective signal dynamics for further application, it would be necessary to know more about their causes.

Early peaks in pupil size changes have already been reported: The decision to select a target (or information leading to this decision) is reflected by pupil size changes occurring even before looking at the target. Indeed, the pupil dynamics observed in our current study correspond to pupil size changes reported by [7] in a cognitive decision task and also to diameter changes reported when spotting an intended objects during visual

search [22, 23]. However, in addition, a subsequent constriction was observed. This constriction could also carry information on the certainty of the foregoing decision and could be included in more sophisticated future selection criteria [29].

Nevertheless, it is still unclear under which circumstances such changes in pupil diameter linked to mental decision processes can be observed. Whether the local minimum may be attributed to constrictions associated with fixations [30], or to decisional processes described by [7] or certainty of choice [29], has to be examined in further studies. For now, these observed signal dynamics provide a promising way to improve human-computer interaction using pupil size changes.

The present data suggest that respective processes are inherent in search and select tasks. For further application, it would be important to know whether similar signal changes could also be obtained when subjects face unintended objects. For example, [31] report considerably larger pupil sizes for targets envisaged to be selected compared to rejected objects; this supports the assumption that pupil dilations selectively accompany intended behavior. However, due to the variety of factors that influence pupil dynamics it is difficult to assign specific characteristics in the signal course to definable cognitive processes. Taken together, although still unclear, early effects regarding pupil size changes which are associated with intentions support our assumption that pupil-size changes can be applicable for active human-computer interaction. The window lengths and thresholds of the pupil-based selection criteria might be modified based on the presented data, e.g. by customizing window length, pupil-based selection times could show less variance than in this experiment; performance in further investigations could be improved.

The current results show for the first time that an implicit physiological event (pupil size changes) can successfully support active target selection in an adequate speed-accuracy tradeoff. Selection times are consistently below one second with accuracy rates being comparable or even better than common approaches [16, 17]. But, in this first approach, PATS did not outperform the benchmarking dwell-time. Including pupil size might improve common ways of computer input, like it has been demonstrated for gaze [32]. Further research has to demonstrate to which extent key usability parameters and the observed pupil dynamics can be replicated and generalized onto different environments and frameworks. Apart from gaze, PATS may be combined with other additional input modalities. Machine learning techniques fitted to the observed signal properties might be able to detect additional signal characteristics that allow both faster and more reliable processing [24]. Further analyses might also determine signal components that provide a valid distinction between pupil size changes that accompany intended and non-intended behavior patterns. Similar to the already existing purely pupil-based HCI input mechanisms, we show that pupil size is a useful variable in order to improve target selection.

6 Conclusion

The integrative interaction concept PATS enriches the established gaze-based target selection procedure with implicit information on pupil diameter. PATS although not outperforming a standard dwell-time approach, was shown to produce only slightly

worse target selection performances with regard to effectivity, efficiency, and usability. A detailed analysis of involved pupil dynamics reveals that early pupil size changes accompany selections in a search and select task. Utilizing these early components in an integrative approach suggest that selection times might be reducible to substantially under 1 s. Summing up, pupil size changes underline the potential of implicit signals even for active target selection in HCI.

7 Acknowledgements

We thank all the volunteers. We would also like to thank Teresa Hirzle in particular for her extraordinary valuable technical assistance and discussions. This study was supported by the SFB 62 by the Deutsche Forschungsgemeinschaft (DFG).

8 References

1. Jacucci, G., Fairclough, S., Solovey, E.T.: Physiological computing. Computer 48, 12–16 (2015)
2. Byrne, E.A., Parasuraman, R.: Psychophysiology and adaptive automation. Biological Psychology 42, 249–268 (1996)
3. Allanson, J., Fairclough, S.H.: A research agenda for physiological computing. Interacting with Computers 16, 857–878 (2004)
4. Ehlers, J., Strauch, C., Georgi, J., Huckauf, A.: Pupil size changes as an active information channel for biofeedback applications. Applied Psychophysiology and Biofeedback, 1–9 (2016)
5. Partala, T., Surakka, V.: Pupil size variation as an indication of affective processing. International Journal of Human-Computer Studies 59, 185–198 (2003)
6. Richer, F., Beatty, J.: Pupillary dilations in movement preparation and execution. Psychophysiology 22, 204–207 (1985)
7. de Gee, Jan Willem, Knapen, T., Donner, T.H.: Decision-related pupil dilation reflects upcoming choice and individual bias. Proceedings of the National Academy of Sciences 111, E618-E625 (2014)
8. Beatty, J., Lucero-Wagoner, B.: The pupillary system. Handbook of Psychophysiology 2, 142–162 (2000)
9. Hess, E.H., Polt, J.M.: Pupil size in relation to mental activity during simple problem-solving. Science 143, 1190–1192 (1964)
10. Wierda, S.M., van Rijn, H., Taatgen, N.A., Martens, S.: Pupil dilation deconvolution reveals the dynamics of attention at high temporal resolution. Proceedings of the National Academy of Sciences 109, 8456–8460 (2012)
11. Simpson, H.M., Hale, S.M.: Pupillary changes during a decision-making task. Perceptual and Motor Skills 29, 495–498 (1969)
12. Slanzi, G., Balazs, J.A., Velásquez, J.D.: Combining eye tracking, pupil dilation and EEG analysis for predicting web users click intention. Information Fusion 35, 51–57 (2017)
13. Pittino, F., Mai, S., Huckauf, A., Pollatos, O.: Effects of hunger on sympathetic activation and attentional processes for physiological computing. In: Proceedings of the 3rd International Conference on Physiological Computing, pp. 542–546 (2016)

14. Jang, Y.-M., Mallipeddi, R., Lee, S., Kwak, H.-W., Lee, M.: Human intention recognition based on eyeball movement pattern and pupil size variation. Neurocomputing 128, 421–432 (2014)
15. Pomplun, M., Sunkara, S.: Pupil dilation as an indicator of cognitive workload in human-computer interaction. In: Proceedings of the International Conference on HCI (2003)
16. Mathôt, S., Melmi, J.-B., van der Linden, L., van der Stigchel, S.: The mind-writing pupil: A human-computer interface based on decoding of covert attention through pupillometry. PloS one 11, e0148805 (2016)
17. Stoll, J., Chatelle, C., Carter, O., Koch, C., Laureys, S., Einhäuser, W.: Pupil responses allow communication in locked-in syndrome patients. Current Biology 23, R647-R648 (2013)
18. Ehlers, J., Strauch, C., Huckauf, A.: A View to a Click: Pupil Size Changes as Input Command in Eyes-only Human-Computer Interaction. International Journal of Human-Computer Studies (submitted)
19. Ekman, I., Poikola, A.W., Mäkäräinen., M.K.: Invisible eni: Using gaze and pupil size to control a game. In: Proceedings of the 2008 CHI Conference on Human Factors in Computing Systems. ACM (2008)
20. Ekman, I., Poikola, A., Mäkäräinen, M., Takala, T., Hämäläinen, P.: Voluntary pupil size change as control in eyes only interaction. In: Proceedings of the symposium on eye tracking research and applications ETRA, pp. 115–118 (2008)
21. Einhäuser, W., Stout, J., Koch, C., Carter, O.: Pupil dilation reflects perceptual selection and predicts subsequent stability in perceptual rivalry. Proceedings of the National Academy of Sciences 105, 1704–1709 (2008)
22. Klingner, J.: Fixation-aligned pupillary response averaging. In: Proceedings of the 2010 Symposium on Eye-Tracking Research & Applications, pp. 275–282. ACM (2010)
23. Privitera, C.M., Renninger, L.W., Carney, T., Klein, S., Aguilar, M.: Pupil dilation during visual target detection. Journal of Vision 10, 3 (2010)
24. Bednarik, R., Vrzakova, H., Hradis, M.: What do you want to do next: a novel approach for intent prediction in gaze-based interaction. In: Proceedings of the symposium on eye tracking research and applications, pp. 83–90. ACM (2012)
25. Majaranta, P., Ahola, U.K., Spakov, O.: Fast eye-typing with an adjustable dwell time. In: Proceedings of the 2009 CHI Conference on Human Factors in Computing Systems. ACM (2009)
26. Lewis, J.R., Sauro, J.: The factor structure of the system usability scale. In: International Conference on Human Centered Design (2009)
27. Peirce, J.W.: PsychoPy—psychophysics software in Python. Journal of Neuroscience Methods 162, 8–13 (2007)
28. Sauro, J., Lewis, J.R.: Correlations among prototypical usability metrics: evidence for the construct of usability. In: Proceedings of the 2009 CHI Conference on Human Factors in Computing Systems, pp. 1609–1618. ACM (2009)
29. Urai, A.E., Braun, A., Donner, T.H.: Pupil-linked arousal is driven by decision uncertainty and alters serial choice bias. Nature Communications 8 (2017)
30. Mathôt, S., Melmi, J.-B., Castet, E.: Intrasaccadic perception triggers pupillary constriction. PeerJ 3, e1150 (2015)
31. Jadue, J., Slanzi, G., Vel, J.: Web user click intention prediction by using pupil dilation analysis. In: Proceedings of the 2015 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT), pp. 433–436 (2015)
32. Zhai, S.: What's in the eyes for attentive input. Communications of the ACM 46, 34–39 (2003)