EyeFrame: real-time memory aid improves human multitasking via domain-general eye tracking procedures

P. Taylor[1,2]*, Noah Bilgrien[1], Ze He[1,3] and Hava T. Siegelmann[1,2]

[1] College of Information and Computer Sciences, University of Massachusetts, Amherst, MA, USA, [2] Neuroscience and Behavior Program, University of Massachusetts, Amherst, MA, USA, [3] Department of Mechanical and Industrial Engineering, University of Massachusetts, Amherst, MA, USA

Objective: We developed an extensively general closed-loop system to improve human interaction in various multitasking scenarios, with semi-autonomous agents, processes, and robots.

Background: Much technology is converging toward semi-independent processes with intermittent human supervision distributed over multiple computerized agents. Human operators multitask notoriously poorly, in part due to cognitive load and limited working memory. To multitask optimally, users must remember task order, e.g., the most neglected task, since longer times not monitoring an element indicates greater probability of need for user input. The secondary task of monitoring attention history over multiple spatial tasks requires similar cognitive resources as primary tasks themselves. Humans can not reliably make more than ~2 decisions/s.

Methods: Participants managed a range of 4–10 semi-autonomous agents performing rescue tasks. To optimize monitoring and controlling multiple agents, we created an automated short-term memory aid, providing visual cues from users' gaze history. Cues indicated when and where to look next, and were derived from an inverse of eye fixation recency.

Results: Contingent eye tracking algorithms drastically improved operator performance, increasing multitasking capacity. The gaze aid reduced biases, and reduced cognitive load, measured by smaller pupil dilation.

Conclusion: Our eye aid likely helped by delegating short-term memory to the computer, and by reducing decision-making load. Past studies used eye position for gaze-aware control and interactive updating of displays in application-specific scenarios, but ours is the first to successfully implement domain-general algorithms. Procedures should generalize well to process control, factory operations, robot control, surveillance, aviation, air traffic control, driving, military, mobile search and rescue, and many tasks where probability of utility is predicted by duration since last attention to a task.

Keywords: eye tracking, gaze tracking, automation, cognitive load, pupil, gaze-aware, human–computer interaction, human–robot interaction

1. Introduction

Interactions with partially autonomous processes are becoming an integral part of human industrial and civil function. Given semi-autonomy, many such tasks can often be monitored by a user at one time. In the course of operating a computer or vehicle, a single human might manage multiple processes, e.g., search and rescue type mobile robots, performing medical supply distribution, patient checkup, general cleanup, firefighting tasks, as well as process control with many dials or readings, security or surveillance monitoring, or other forms of human-based monitoring or tracking tasks. Generally, each automated agent or process only needs intermittent supervision and guidance from a human to optimize performance, and thus a single user can remotely operate or supervise multiple entities, for efficiency of labor. When controlling multiple automated processes at once, the user must decide how to distribute attention across each task. Even if the operator conducts the same type of task with each automated process, this form of human–system interaction requires a multitasking effort.

Unfortunately, most people are notoriously poor multitaskers (Watson and Strayer, 2010) and can remain unaware of visually subtle cues that indicate the need for user input. Further complicating the situation, individuals who perform worst at multitasking actually perceive they are better at multitasking, demonstrated by negative correlations between ability and perception of ability in large studies (Sanbonmatsu et al., 2013). To make matters worse, humans often naturally develop a plethora of biases of attention and perception. To address many of these issues, divided attention performance has been studied for many years (Kahneman, 1973; Gopher, 1993). A further difficulty in multitasking is that brains rely heavily upon prediction and, fundamentally, are incapable of knowing what important information they have missed.

Eye tracking to ascertain point of gaze is a highly effective method of determining where people orient their attention (Just and Carpenter, 1976; Nielsen and Pernice, 2010), as well as what they deem important (Buswell, 1935; Yarbus, 1967). Traditionally, eye tracking informed post-experiment analysis, rather than helping users in the field in real-time. For example, a study might analyze optimal gaze strategies in high-performing groups, and then at a later date, train new users on those previously discovered optimal search strategies (Rosch and Vogel-Walcutt, 2013). For example, studies have trained novice drivers' gaze to mimic experienced drivers with lower crash risk (Taylor et al., 2013).

Alternatively, eye movement strategies can be employed to optimize real-time task performance, since many eye-movements are capable of being intentionally controlled. For those eye movements that cannot easily be intentionally controlled, salient "pop-out" cues (e.g., flashing red box around target) can reliably direct attention in a more automatic, bottom-up manner. As we discuss further, many eye tracking systems have been developed for real-time control, with very few attempting pure assistance, though none were both successful and domain-general (Taylor et al., 2015). There appears to be a need for such an assistive system. Here, we tested a solution, which was uniquely domain-general, non-interfering, purely gaze-aware, and most importantly, yielded large benefits in performance.

The work described here maintained several hypotheses and operational predictions. When a user manages multiple tasks in the real world, the primary task must be performed, while a secondary task of correctly remembering their attention history also often contributes to optimizing performance. Here, we provided participants a computerized system to perform the secondary task of remembering, processing, and then displaying actions derived from gaze history. This treatment was predicted to increase users' ability to perform any task requiring semi-random gaze patterns over a large mufti-faceted display or set of dials and readouts. Supplying participants with this highlighted inverse of eye gaze recency may improve user performance by delegating short-term working memory for attention history to the computer. This could generalize to many multitasking scenarios where probability of needed action on a given single task increases over duration since last interaction with the task.
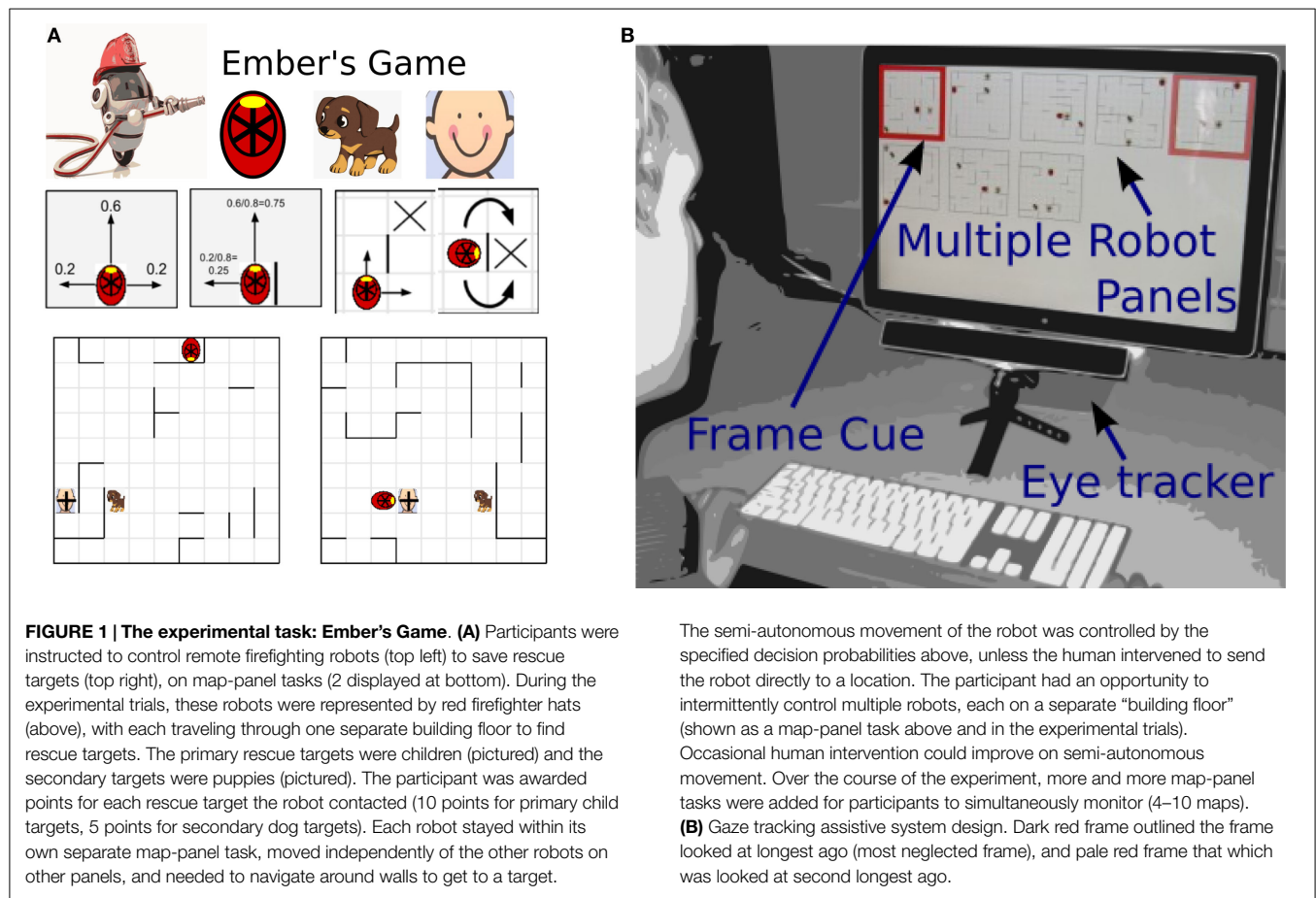
2. Materials and Methods

We designed an eye tracking system to assist users in a monitoring and search task with multiple simulated agents. Assistance was specified by an algorithm to highlight in real-time the most neglected task elements. The game-like task designed for participants required human monitoring and interaction with multiple rescue robots at one time. Our algorithm helped users determine where to look, to improve performance and reduce cognitive load. Our goal was to create a game task, which tested more a general utility of design, to extrapolate to many future tasks, in which the probability of need per task increases with time.

2.1. Experimental Task for Human Subjects: Ember's Game

To participants, the experimental task was presented as "Ember's Game." It was used to test our short-term working memory aid. A user managed from 4 to 10 mobile virtual agents (robots) on a computer monitor at one time, in sequential experimental blocks (trials). The game provided a measurable means to assess visual-spatial multitasking. Beyond the experimental questions it asked, the simulation was designed to be engaging for the user, practically oriented, and to represent a broad variety of human–robot, human–computer, or process-control tasks encountered in the fields of human factors and engineering psychology for industrial, civil, or military applications.

In each task, a human operator directed a firefighting robot to move semi-autonomously through a natural building-like environment, to rescue people and pets from a fire elsewhere in the building (**Figure 1**). Three types of objects were included in the game: (1) firefighting robots, which autonomously explored the space via a semi-random walk, saving firefighting victims upon contact, (2) primary targets, which were human rescue victims, when saved earned the operator 10 points, and (3) secondary targets, which were puppies, earned the operator 5 points when saved. The instructed goal of the game was to obtain as many points as possible by saving rescue victims within a fixed time limit. The user could improve on the robot's performance by assuming control of the semi-random walk behavior to send the robot more directly to targets. This human intervention

FIGURE 1 | The experimental task: Ember's Game. (A) Participants were instructed to control remote firefighting robots (top left) to save rescue targets (top right), on map-panel tasks (2 displayed at bottom). During the experimental trials, these robots were represented by red firefighter hats (above), with each traveling through one separate building floor to find rescue targets. The primary rescue targets were children (pictured) and the secondary targets were puppies (pictured). The participant was awarded points for each rescue target the robot contacted (10 points for primary child targets, 5 points for secondary dog targets). Each robot stayed within its own separate map-panel task, moved independently of the other robots on other panels, and needed to navigate around walls to get to a target.

The semi-autonomous movement of the robot was controlled by the specified decision probabilities above, unless the human intervened to send the robot directly to a location. The participant had an opportunity to intermittently control multiple robots, each on a separate "building floor" (shown as a map-panel task above and in the experimental trials). Occasional human intervention could improve on semi-autonomous movement. Over the course of the experiment, more and more map-panel tasks were added for participants to simultaneously monitor (4–10 maps). (B) Gaze tracking assistive system design. Dark red frame outlined the frame looked at longest ago (most neglected frame), and pale red frame that which was looked at second longest ago.

would accelerate rescue times over the less targeted random walk movement.

Participants were instructed to manage multiple independent sections of the burning building at one time, where each building section had its own maze, firefighter, and rescue targets (**Figure 1**). The set of building map-panel tasks were arrayed across a computer screen in a grid formation (**Figure 1**). Since the robots were semi-autonomous and could follow human instructions, yet take around 1–10 s to reach their targets, each map-panel task only required user interaction intermittently. When a user interacted with a building map-panel task and satisfied task requirements ("saves" a rescue target), new rescue targets could appear afterwards, which was an opportunity to gain more points. Thus, the probability of these points having re-appeared on that map panel increased with time. Consequently, the user could optimize their own behavior by switching between these tasks to interact with the map-panel task, which required intervention. Each user played multiple experimental blocks of the experiment, first with 4 map-panel tasks, then 5, with up to 10 map tasks, requiring optimal switching between map panels to score well.

While Ember's game is simple, it particularly taxes spatial working memory used in visual screen monitoring environments and should generalize well to many tasks in which the probability of task utility increases as a function of time since user intervention with the given task element. These methods are potentially beneficial, in part because they may yield improvement

via a simple visual reminder for many types of working memory intensive spatial monitoring tasks.

2.2. Manipulated Independent Variable: Cue-Type

We employed eye tracking to produce contingent highlighting via gaze history across the set of all maps (**Figure 1**). This gaze history was inverted compared to traditional heat-maps for gaze history; in this study, the most neglected map-panel tasks were highlighted for the user. In this gaze-aware display, elements in the display array were highlighted in color-steps, like a gradient of time since the user last looked at the display, with the elements looked at longest ago highlighted most saliently. This gradient estimated probabilities of utility, which notably, need not be independently discoverable by the computer (e.g., via computer vision).

The experiment was a between-subjects design with three groups of participants (one test and two control), determined by three "frame types" that were employed: (1) helpful frame cues, (2) randomly moving frames ("active" control), and (3) no frames ("absent" control). These three experimental groups of participants are detailed further here:

1. Helpful frame cues ("On" experimental condition): our short-term working memory aid algorithm drew a red frame cue around the most neglected map-panel task looked at longest ago, which may thus have been most in need of the user's attention, and drew a pale red line around the map-panel

task looked at the second longest ago in time. To do so, every monitor-refresh, the eye tracker notified the computer and game of the location of gaze. Frame cues for the most neglected map panels were removed and re-updated if the user glanced at or interacted with a map panel. As a backup for eye tracker error, if the user clicked on a map-panel task, the frame was removed, since the user had manually interacted with that map-panel task. In summary, a map-panel task glanced at recently was not highlighted, while a neglected map-panel task glanced at longest ago was brightly highlighted. This history of successive gazes was then used to provide real-time visual cues for more effective task switching between maps, throughout the entire course of the game.

2. Randomly moving frame highlighting ("Random" control condition): mimicked the Helpful condition (On) frames in most other aspects, other than their relationship to user gaze. Using the same physical stimulus as the Helpful condition, two randomly chosen map-panel tasks were highlighted at any given time, and stayed highlighted for a random amount of time between 1 and 2 s, closely approximating the amount of time a frame stays highlighted when a user looks at a given map-panel task. Random highlighting helped to control for novelty or pop-out effects using the same physical stimulus as during Helpful frames, but without the helpful information. Users were notified before they started playing that the red frames were random and irrelevant to the game, and the user was able to choose their own strategy for switching between maps.

3. No map-panel task frame highlighting ("Off" control condition): was employed in the second control condition, and the user chose their own strategy for switching between map-panel tasks.

Each group experienced the same progression of 4 simultaneous maps to manage, up to 10, creating 6 levels of the map-panel task number factor.

2.3. Experimental Procedure

The experimental sequence was as follows: before starting, each participant received and signed an informed consent. Participants were randomly assigned to conditions (Helpful frame cues, Random frames, Off frames). The eye tracker apparatus was explained to participants. Where possible, all experimental groups received identical instructions (except as required for condition-differences) for playing Ember's game, and any questions were answered (instructions included in Supplementary Material). Before starting, lights were dimmed to a level consistent across participants, for calibration reliability and pupil dilation consistency. Participants were given practice playing, first training 1 building map-panel task at a time, then training with 2 maps at the same time, followed by 3 maps, with 60 s for each block. Then, experimental trials began (different highlighting for each group), with participants playing an experimental block containing 4 map-panel tasks simultaneously, and subsequent blocks containing 5 through 10 maps simultaneously (150 s each block), adding 1 map-panel task per block.

2.4. Measurement and Dependent Variables

A variety of measures were recorded to index working memory load, cognitive load, and distribution of humans' visual attention. Behavioral performance and reaction times were measured to analyze strategies. Point of gaze was recorded throughout the task. Measures of pupil dilation indexed cognitive load. Many studies have shown that pupil dilation is a reliable measure of cognitive load under certain conditions (Hess and Polt, 1964; Kahneman and Beatty, 1967), with more mental effort typically assumed to be associated with larger pupil size; expanded in the Discussion.

Data logging included: the status of all experimental variables on every refresh (at 30 Hz) during experimental trials. Behavioral data were indexed by location and status of all game elements, such as robot location, path location, target location, and time of target detection. Eye data were indexed by left and right point of gaze on the screen (x, y coordinates) at the refresh rate frequency, the calibration quality data (error quantity) before every new block, and pupil dilation of left and right eye diameter in millimeters at every time-step. Mouse location (x, y coordinates) was recorded at the same frequency for comparison to gaze data.

2.5. Participants

A total of 44 human subjects participated in Ember's game. All procedures complied with departmental and university guidelines for research with human participants and were approved by the university institutional review board. Participants were compensated for their time with $5 USD. Data were not excluded based on behavioral task performance, in order to obtain a generalizable sample of individual variation on performance of the task, while avoiding a restriction of range (Myers et al., 2010). Two participants with vision correction causing poor calibration quality for entire blocks were excluded, leaving 42 subjects. No data were excluded within this pool of subjects. Minimal pilot data were collected using extended game-play and testing on the experimenters themselves, though these data were not included due to their short length and differing parameters from the experimental subjects. Between-subjects design was selected, to avoid "order" and "training" effects which are present in within-subjects designs, particularly with 2 of 3 groups being controls. Each participant reported past video game experience, current vision correction, age, and sleep measures for the previous several days; this was done after rather than before experimental task participation to prevent bias.

2.6. Technical Implementation

We used a desk-mounted GazePoint GP3 eye tracker to pinpoint the users' point of gaze, i.e., the point on the screen the user is fixating. This eye tracker has an accuracy of [0.5–1] degrees of visual angle and 60 Hz update rate. Nine-point calibration was performed immediately before every 150 s new block. Python and PyGame were used to program the experiment, and interfaced with the eye tracker's open standard API via TCP/IP, generously provided by GazePoint (http://www.gazept.com). Gaze on a panel (for the Helpful frames condition) was inferred if the user had been looking at a map-panel task continuously for 10 frames (about 1/3 of a second). Fixations are generally considered to be

roughly 80–100 ms, and this 300 ms duration was chosen because it approximated the amount of time a player had to look at a map-panel task to focus on it and obtain the information they need, while also being long enough to significantly reduce the effect of eye tracker error.

2.7. Statistical Analysis

Most statistics are displayed within figures themselves, either (1) as standard error of the mean (SEM) bars, which, in our experiment, conservatively indicate statistically significant differences between groups by approximating t-tests if SEM bars are not overlapping between conditions, as explained below, (2) as pairwise t-tests superimposed on map bias task arrays, (3) as Pearson's product moment correlation coefficient r and p-values superimposed on scatter-plots, and (4) as effect sizes calculated via Cohen's d (Table S1 in Supplementary Material).

The t-statistic is defined as the difference between the means of two compared groups, divided by the SEM, $(u_1 - u_2)/SEM$. Thus, within the parameters of this experiment (and above any typical n) it is a mathematical necessity that when the SEM bars do not overlap, a t-test on those same data would be significant above an alpha criterion of around $p < 0.03$ for a one-tailed t-test for effects in the expected direction (as they were this experiment).

The low number of tests within proposed statistical families, the presence of consistent global trends, and guidelines discussed below, all argue against correcting p-values themselves for multiple comparisons (Rothman, 1990; Saville, 1990; Perneger, 1998; Feise, 2002; Gelman et al., 2012). Further, many statisticians do not recommend numerically correcting for multiple comparisons (Rothman, 1990; Saville, 1990). Rather, it is often recommended to document individual uncorrected p-values, while being transparent that no correction was performed for multiple comparisons. In light of the differing opinions of what defines a statistical family, we provided Bonferroni-corrected alpha thresholds, though these are known to be overly conservative (Perneger, 1998), especially for measures predicted to be correlated, as ours were. The

p-values presented for these experiments in Table S2 in Supplementary Material should be interpreted while considering that alpha thresholds would traditionally be significant at $p < 0.05$, while with seven hypothetical measures in a family the Bonferroni threshold was adjusted to around $p < 0.007$, with 21 hypothetical tests per family measure the Bonferroni alpha would be adjusted to approximately $p < 0.002$, or with an evaluation more similar to Dunnett's method, a corrected measure for 10 tests was around $p < 0.004$. When these corrections were applied to the p-values found in Table S2 in Supplementary Material, no conclusions were changed. *Post hoc*, the percentage of significant tests out of the set of all tests can be observed, and evaluated as to whether it deviates or conforms to the statistically expected percentage. This is similar to methods, such as Holm's, which rank p-values, or methods which evaluate the percentage of significant tests out of the set of total tests, and it should be noted that our conclusions rested not upon a single test, but upon globally uniform patterns.
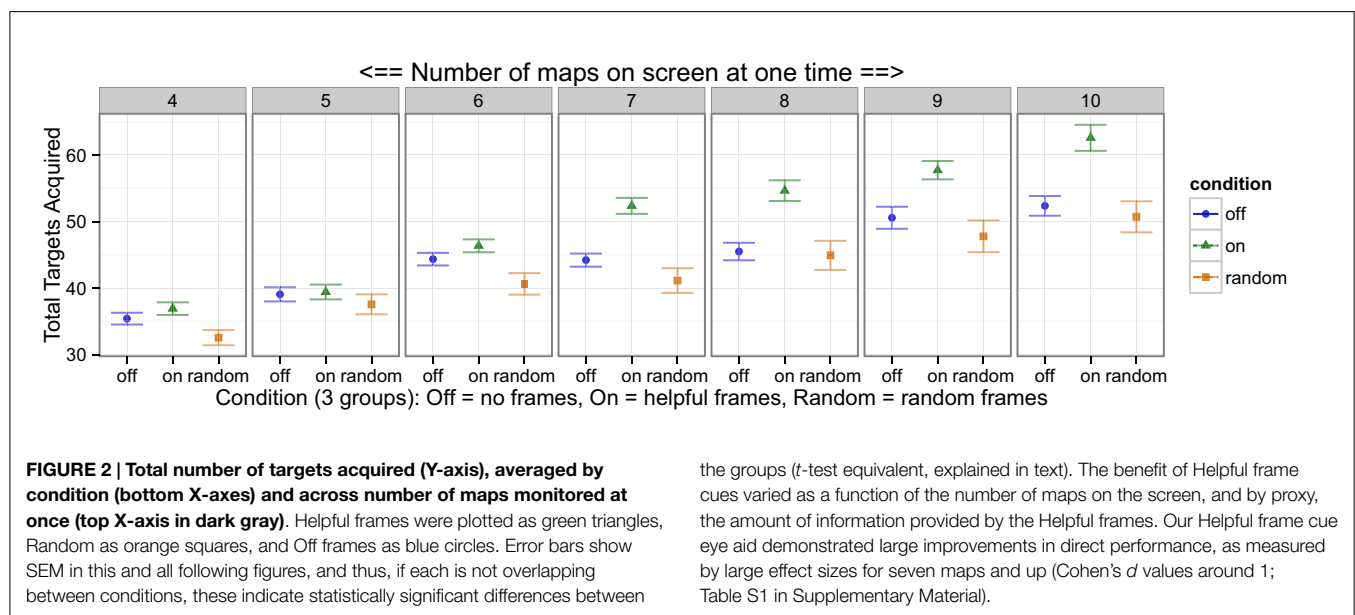
Data processing and plotting were programed in the R-project statistical environment (Core Team, 2013).

3. Results

Below we detailed the practically relevant results with bearing for applications in human factors and related fields. For a cognitive and mechanistic deconstruction of the effect of our cuing system on human participants, we provided an additional elaborate analysis in Taylor et al. (2015).

3.1. Helpful Frame Cues Improved Performance

Most importantly, our results demonstrated large performance improvements in mean game scores for operators using our memory aid, as compared to the two control groups (**Figure 2**). The Random frames control condition showed slightly reduced scores as compared to the Off frames control condition, indicating that the randomly moving frames may have been distracting

FIGURE 2 | Total number of targets acquired (Y-axis), averaged by condition (bottom X-axes) and across number of maps monitored at once (top X-axis in dark gray). Helpful frames were plotted as green triangles, Random as orange squares, and Off frames as blue circles. Error bars show SEM in this and all following figures, and thus, if each is not overlapping between conditions, these indicate statistically significant differences between

the groups (t-test equivalent, explained in text). The benefit of Helpful frame cues varied as a function of the number of maps on the screen, and by proxy, the amount of information provided by the Helpful frames. Our Helpful frame cue eye aid demonstrated large improvements in direct performance, as measured by large effect sizes for seven maps and up (Cohen's d values around 1; Table S1 in Supplementary Material).

to the users, and validating the importance of a second control. The benefit of the eye tracker appears largest for larger numbers of maps, likely because the eye tracking system compensated for the inability to optimize switching across seven or more panels.
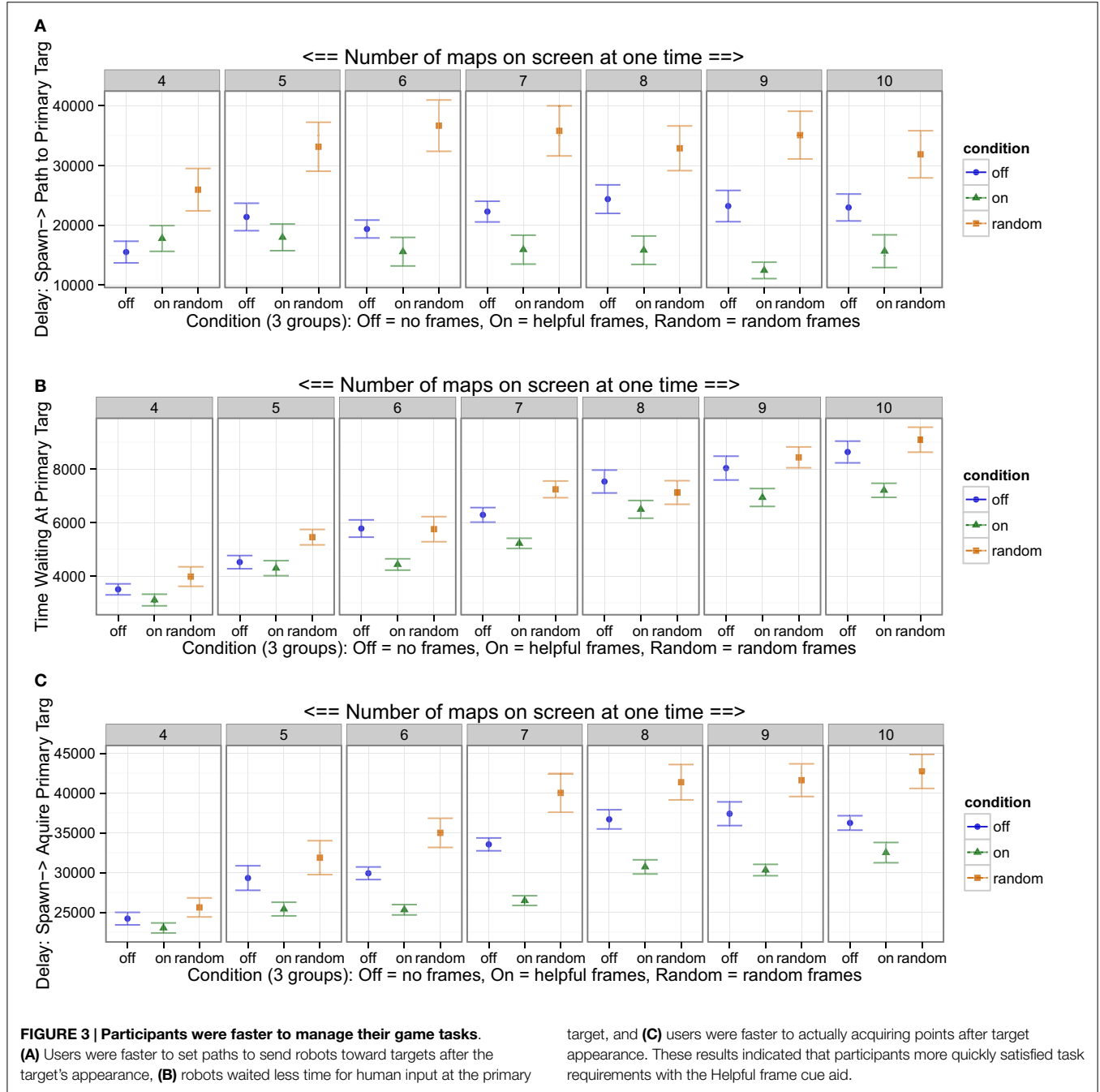
3.2. Reaction Times were Improved

As would be predicted by higher scores, participants were also faster at managing their task, measured in several ways. First, reaction times were faster to targets, as measured by the delay from a target appearing, to the user directing the robot toward it (**Figure 3A**). Second, participants were faster to assist waiting

robots, which required user input to actually save the primary targets (**Figure 3B**). Third, the delay from the time a target spawned until the target was actually acquired for points was shorter in the Helpful frame cue condition (**Figure 3C**). Interestingly, though total scores in the Off frames control were only marginally better than Random frames control (above), reaction times were reliably worse for the Random condition than the Off frames condition; compare **Figure 2** with **Figure 3**.
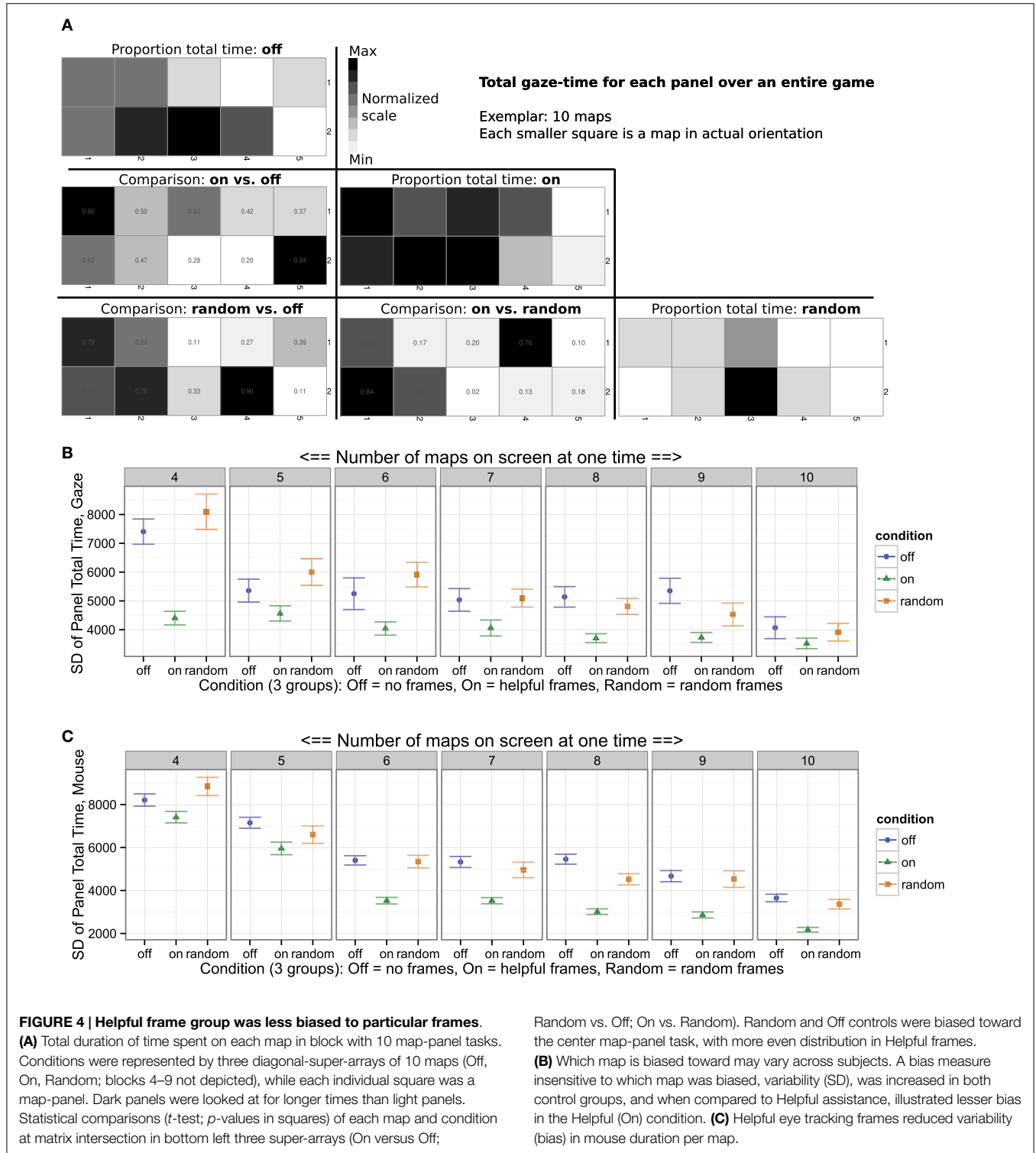
3.3. Global Bias was Reduced

When interacting with this simulation, like many real-world tasks, it is often important to eliminate biases when these biases are

FIGURE 3 | Participants were faster to manage their game tasks. **(A)** Users were faster to set paths to send robots toward targets after the target's appearance, **(B)** robots waited less time for human input at the primary

target, and **(C)** users were faster to actually acquiring points after target appearance. These results indicated that participants more quickly satisfied task requirements with the Helpful frame cue aid.

either artifactual or irrational. In this case, a bias would take the form of greater time spent attending to a single map-panel task, over the time spent looking at equally relevant other map-panel tasks on the screen at the same time. We employed two possible measures of bias. The first measure averaged the duration of time spent on each map-panel task, with a set of averages for each map-panel task in each experimental block (each experimental

block has a different number of map-panel tasks). This first measure took the form of a heat overlay displaying the cumulative time the gaze spent on each map-panel task during the entire block (**Figure 4A**). In this exemplar case of 10 map-panel tasks, most subjects in Random and Off conditions were biased toward the maps in the middle of the screen (darker colors in the middle squares), while attention was more evenly spread in the Helpful

FIGURE 4 | Helpful frame group was less biased to particular frames.
(A) Total duration of time spent on each map in block with 10 map-panel tasks. Conditions were represented by three diagonal-super-arrays of 10 maps (Off, On, Random; blocks 4–9 not depicted), while each individual square was a map-panel. Dark panels were looked at for longer times than light panels. Statistical comparisons (t-test; p-values in squares) of each map and condition at matrix intersection in bottom left three super-arrays (On versus Off;

Random vs. Off; On vs. Random). Random and Off controls were biased toward the center map-panel task, with more even distribution in Helpful frames.
(B) Which map is biased toward may vary across subjects. A bias measure insensitive to which map was biased, variability (SD), was increased in both control groups, and when compared to Helpful assistance, illustrated lesser bias in the Helpful (On) condition. **(C)** Helpful eye tracking frames reduced variability (bias) in mouse duration per map.

frame experimental condition (more evenly spread heat map over the whole array). This measure agreed with previous literature that human subjects have an "edge effect," being biased toward the center (Parasuraman, 1986). One limitation of this grand-mean measure is that variability is primarily sensitive to spatial biases which occur across all subjects.

Thus, to systematically quantify any biases at an individual subject level, we calculated the standard deviation (SD) of total times on each panel, across all map-panel tasks. This described the variation in cumulative time the eyes spent on each of the maps in an experimental block and map number condition (**Figure 4B**). For example, in the seven maps Helpful condition, the set of cumulative times across each of the ten maps had a low variability and thus each map-panel task was looked at for a similar amount of time, while in the control conditions the set of cumulative times across each of the ten maps had a higher variability indicating that there was greater bias toward some maps away from others. Confirming the eye gaze results, a similar reduction in bias was observed for the time the mouse spent over each map-panel task (**Figure 4C**). This decreased variability of total time spent on any single map-panel task compared to the rest indicated improved consistency of time spent on each map-panel task with the eye tracker, while participants favored some map-panel tasks (robots) irrationally when they did not have the eye tracker's help.

It is notable that not all bias is bad, since some task elements may be more important or require more frequent input than others. Current experiments in our lab automated rational biases by unevenly weighting delay for elements of the array panel, providing a gaze-contingent system to distribute gazes accordingly.

3.4. Measures of Cognitive Load were Reduced

To assess measures of cognitive load, pupil dilation was recorded, both over the course of each block and averaged across blocks. We demonstrated the dynamics of pupil dilation over the course of a trial, using data from the "7 map-panel tasks" block as an exemplar, since this is where behavioral benefits started to convincingly appear. To do so, three time-traces for pupil dilation for each condition were plotted over time (**Figure 5A**). For 7 map-panel tasks, the Helpful condition had the lowest pupil dilation relative to the other conditions, while Off frames was in the middle, and Random frames had the largest pupil dilation. Pupil dilation was then collapsed over the entire trial, for each condition and number of maps. Pupil dilation was reliably smaller with Helpful frames than in the Random frames condition (**Figure 5B**); though not predicted, pupil dilation did not significantly differ between the Off frames and Helpful frames condition. This finding confirmed the benefit of including a second control group using the same physical condition with no information (Random frames). It is possible that the Random frames were distracting, and took cognitive effort to ignore, compared to the Helpful frames. Variations in pupil dilation over map number could be generally explained by experience, novelty, fatigue, or training, though since this factor was not experimentally manipulated to explore the effect of number of items on the screen, reliable interpretations of map number effects can not be drawn. Interestingly, when pairing pupil dilation and score for each individual subject, within the Helpful condition,

larger pupils were associated with higher total scores (**Figure 5C**). These results suggest operators exerting more "effort" had larger pupils.

3.5. Participant Sample Statistics

Lastly, we thoroughly confirmed there were no incidental differences between subject groups in each condition for features known to influence experimental performance. To do so, we tested the null hypothesis that each group had the same population mean using ANOVA for the following measures: (A) hours of sleep in the previous week did not differ ($F = 0.2$, $p < 0.8$), (B) age in years (mean $= 26$) did not differ ($F = 1.2$, $p < 0.3$), and (C) multiple measures of video game experience did not vary between conditions, as measured by post-experimental surveys assessing multiple measures of gaming frequency ($F = 0.8$, $p < 0.5$ – days/year; $F = 0.8$, $p < 0.5$ – h/week), and gaming history ($F = 0.1$, $p < 0.9$ – duration; $F = 0.5$, $p < 0.6$ – age started playing).
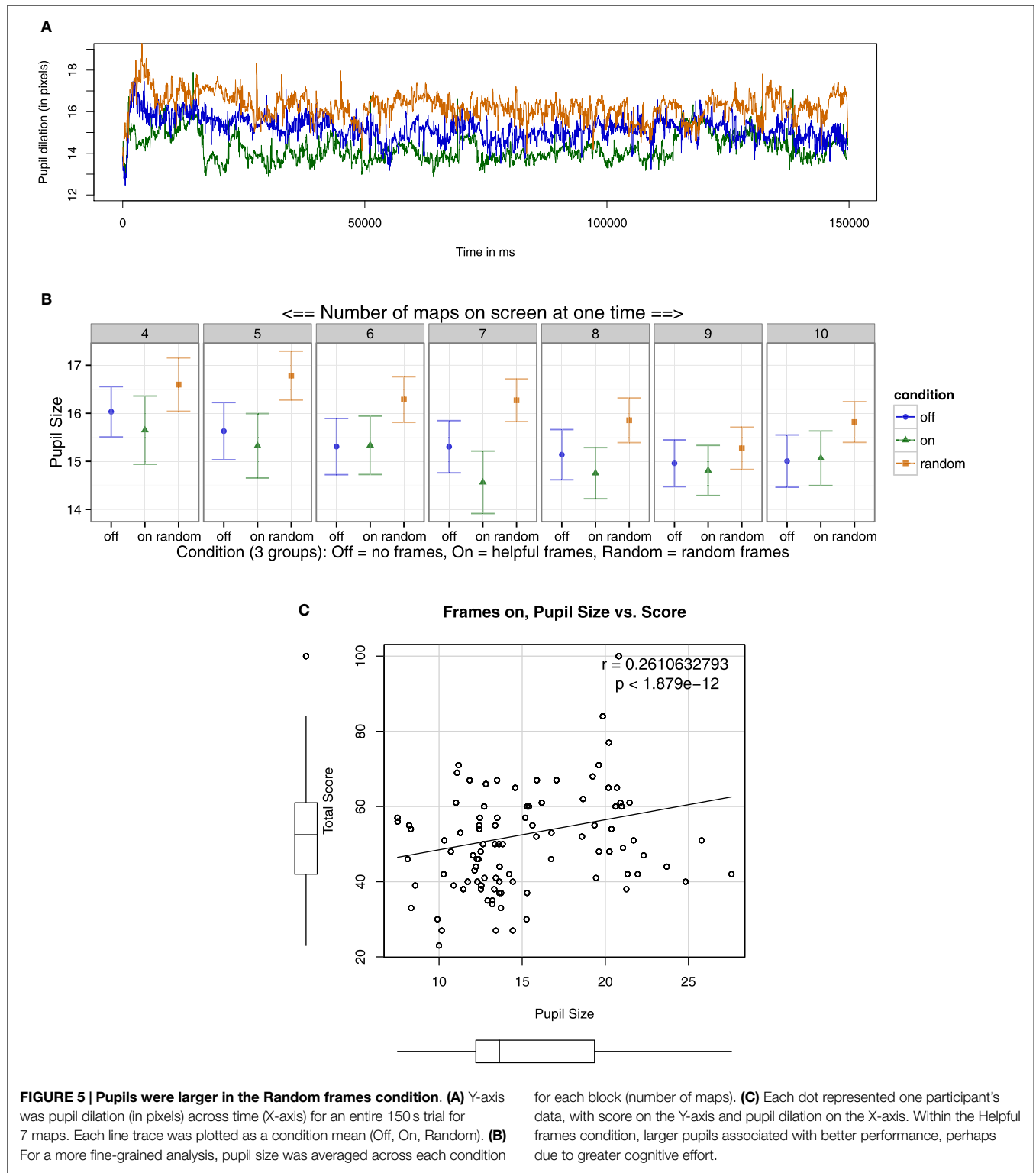
4. Discussion

4.1. Real-World Performance

Our gaze tracking algorithms improved human operator performance, with very large effect sizes, both for pure performance scores, and for reaction times (Cohen's d; Table S1 in Supplementary Material). These improvements have a good chance of being reflected in real-world applications of such an algorithm, because the scenarios tested here were designed to have realistic features and be difficult. For example, humans are reliably worse at dual-target compared to single target searches, even though such search needs arise in every day life and in some work settings, e.g., when scanning x-rays for both explosive devices and metal objects (Menneer et al., 2012). To relate, our task employed such a dual-target search for conjunction targets. While pop out type searches (a stand-out object) are fast and hypothesized to be performed quickly in parallel pre-attentively, conjunction searches (multiple features) are serial and slow (Treisman and Gelade, 1980). Our study included conjunction stimuli that required a more difficult serial search to find. Conversely, the red frame cue acted as a pop out stimulus and was thus very easy to find.

4.2. Bias

There are a myriad of biases which can be adopted by a user, and visual biases are some of the most well known. For example, search tends to be biased toward central regions of available visual space, coined an "edge effect" (Parasuraman, 1986). We observed this type of effect in the two control conditions, and our assistive system greatly reduced this bias (**Figure 4A**). Visual biases, likely derived from reading, have also been observed for the upper left of a display (Megaw and Richardson, 1979), which we may have also observed. Further, individuals may vary in biases for or against different task elements, but these may not be present in group means. To address this issue, we demonstrated that the variability in looking time at each task was reduced in the assistive frames condition compared to either control, confirming more strongly that our system reduced user bias. It should be noted that not all bias is undesirable, and we extended these concepts in the Future Directions Section below.

FIGURE 5 | Pupils were larger in the Random frames condition. **(A)** Y-axis was pupil dilation (in pixels) across time (X-axis) for an entire 150 s trial for 7 maps. Each line trace was plotted as a condition mean (Off, On, Random). **(B)** For a more fine-grained analysis, pupil size was averaged across each condition for each block (number of maps). **(C)** Each dot represented one participant's data, with score on the Y-axis and pupil dilation on the X-axis. Within the Helpful frames condition, larger pupils associated with better performance, perhaps due to greater cognitive effort.

4.3. Task Switching

Human operators appear to be unable to reliably make greater than two decisions per second (Craik, 1948; Elkind and Sprague, 1961; Fitts and Posner, 1967; Debecker and Desmedt, 1970). Our system potentially eliminates one decision per second, or per map switch, a non-trivial benefit. Visual cues for task switching may

assist operators (Allport et al., 1994; Wickens, 1997); this is particularly true when cuing important tasks (Wiener and Curry, 1980; Funk, 1996; Hammer, 1999). However, we are the first to implement such reminders via eye tracking in a manner that can easily be implemented across domains and platforms. In many multitask scenarios, like 8 or more maps here, it is likely optimal to have

an automated task-management strategy. Our algorithm leverages these phenomena to optimize the primary task, while secondary tasks (remembering order) are also beneficial to perform, this can be automated reliably. Our frame cue aid appeared to assist participants to more fully attend to a single map at a time, with efficient task switching between maps. Further, decision-making contributes to cognitive load.

4.4. Cognitive Load

In addition to attention, eye tracking can also be used to asses several nebulous mental states, including that of cognitive effort. For example, over 50 years of research suggests measuring pupil dilation over time (pupillometry) is a reliable measure of cognitive load, where larger pupils indicate greater load or arousal in controlled lighting conditions (Hess and Polt, 1964; Kahneman and Beatty, 1966, 1967; Bradshaw, 1968; Simpson and Hale, 1969; Goldwater, 1972; Hyona et al., 1995; Granholm et al., 1996; Just et al., 2003; Granholm and Steinhauer, 2004; Van Gerven et al., 2004; Haapalainen et al., 2010; Piquado et al., 2010; Laeng et al., 2012; Wierda et al., 2012; Hwang et al., 2013; Zekveld and Kramer, 2014; Zekveld et al., 2014). Our system also appeared to reduce cognitive load, as illustrated by reduced pupil size (lower cognitive effort) in the Helpful condition compared to the Random control (**Figure 5**). These results, in particular, were not unequivocal, though the general trend matched expectations.

4.5. Novelty of the Real-Time Eye Tracking System

Previous work has used point of gaze to update displays in real-time, contingent on pupil size or on location of gaze. Both types are summarized here.

4.5.1. Pupil Size in Real-Time

The DARPA augmented cognition initiative primarily evaluated pupil dilation as a measure of cognitive load, some in real time (Marshall, 2002; Marshall et al., 2003; St John et al., 2003; Taylor et al., 2003; Raley et al., 2004; St. John et al., 2004; Johnson et al., 2005; Mathan et al., 2005; Russo et al., 2005; Ververs et al., 2005; De Greef et al., 2007; Coyne et al., 2009). Though some attempted to use gaze location (Barber et al., 2008), none were successful.

4.5.2. Real-Time Eye Tracking as an Experimental Tool

Contingent eye tracking modifies computer displays in real-time based on gaze location and was traditionally used in psychology experiments, though usually to impede users, not to optimize performance. During the development of eye tracking technology, psycholinguists used methods, such as the moving window paradigm (Reder, 1973; McConkie and Rayner, 1975), which interferes with the participant by turning the upcoming periphery into noise or random stimuli while reading, to reduce parafoveal preview. Many other paradigms have been employed to impede participants, such as the moving mask paradigm (Rayner and Bertera, 1979; Castelhano and Henderson, 2008; Miellet et al., 2010), the parafoveal magnification paradigm

(Miellet et al., 2009), or a central hole (Shimojo et al., 2003). Interfering with performance can successfully function to probe participant's abilities in experiments. However, none attempted to improve performance, instead degrading it.

4.5.3. Real-Time Gaze Control

Gaze-based systems to control computer displays, wheelchairs, or other robots have been extensively developed. These methods often aimed to move cursors, wheelchairs, accessories, graphical menus, zoom of windows, display context-sensitive presentation of information, while also including systems for mouse clicks (Jacob, 1990, 1991, 1993a,b; Jakob, 1998; Zhai et al., 1999; Tanriverdi and Jacob, 2000; Ashmore et al., 2005; Laqua et al., 2007; Liu et al., 2012; Sundstedt, 2012; Hild et al., 2013; Wankhede et al., 2013). Such systems have also been employed for robot and swarm robot control (Carlson and Demiris, 2009; Couture-Beil et al., 2010; Monajjemi et al., 2013). Many of these methods were primarily for control, though could be interpreted to have some assistive component, such as displaying context-sensitive information. Our gaze-aware system can improve performance, without direct input, and may assist operators in a variety of scenarios, both control and non-control. This purely assistive (rather than control) algorithm may complement some of these control systems.

4.5.4. Gaze-Aware Assistive Systems

Gaze-aware systems are both much less common, and the interfaces have been domain specific, such as those used for reading, menu selection, view scrolling, or information display (Bolt, 1981; Starker and Bolt, 1990; Sibert and Jacob, 2000; Hyrskykari et al., 2003; Fono and Vertegaal, 2004, 2005; Iqbal and Bailey, 2004; Ohno, 2004, 2007; Spakov and Miniotas, 2005; Hyrskykari, 2006; Merten and Conati, 2006; Kumar et al., 2007; Buscher et al., 2008; Bulling et al., 2011). Few human–robot studies have taken gaze location into consideration for improving human task performance (DeJong et al., 2011), with such studies also being limited to application-specific industrial goals. Often such systems predict the participants' gaze location in tasks like map scanning, reading, eye-typing, or entertainment media (Goldberg and Schryver, 1993, 1995; Salvucci, 1999; Qvarfordt and Zhai, 2005; Bee et al., 2006; Buscher and Dengel, 2008; Jie and Clark, 2008; Xu et al., 2008; Hwang et al., 2013). Some studies have succeeded at domain-general methods of prediction (Hwang et al., 2013); however, these methods all require advanced computational methods, such as image processing and machine learning. We argue that eliminating prediction and focusing on describing gaze history yields more general utility. Very few studies have attempted to create domain-general assistive systems via real-time eye tracking. However, such attempts have had drawbacks, for example, by obscuring the display permanently (e.g., make everything glanced at obscured), or only applying to search, rather than monitoring or control tasks (Pavel et al., 2003; Roy et al., 2004; Bosse et al., 2007). The system presented here improves upon these previous deficiencies; it is domain-general, easy to use, closed-loop, non-predictive, does not require advanced computational methods, such as computer vision, and most importantly, is effective. Despite the need and benefit from this situation-blind

multitasking aid, and the seeming obviousness of our solution, no domain-general solutions such as this exist to date. The experimental results showed large benefits, and there is potential for wide generalization.

4.6. Behavioral Mechanisms and In-Depth Discussion

We provide an extended in-depth eye tracking analysis of these studies, as well as further discussion of prior assistive systems, divided attention, bias, supervisory sampling, search, task switching, working memory, augmented cognition, contingent eye tracking, gaze-control, gaze-aware displays, and other topics in Taylor et al. (2015).

4.7. Conclusion and Future Directions

Our real-time gaze aid was successfully domain-general in an unprecedented way, succeeding where previous studies fell short, either by being application-specific, control oriented, or by interfering with the user in a non-intuitive manner. It is likely that our algorithm, of displaying the inverse of gaze recency, could benefit many multi-monitoring or tracking tasks where probability of utility relates to duration since last attention to an element. These applications might include process control, multitasking factory operations, multi-robot control, multi-panel surveillance, aviation, air traffic control, driving, military, mobile search and rescue, or others.

For future work, user-assisting systems should be flexible to the task needs, i.e., should be able to choose and change how frequently the cues appear, depending on the frequency with which agents need input. Currently, our lab has developed a platform-independent general overlay application, which includes weighting the value of each task in the array to be looked at. For example, some dials might be more important than others or require more frequent input. Further, to ensure compatibility with a variety of tasks, we incorporated varying colors, levels of visual transparency, and types of cue. Also, user-assisting systems should be practical and able to be implemented broadly, and thus we incorporated the use of the mouse rather than eye in a similar application, for broad use without eye tracking.

Data Sharing

Source code for the Ember's game experiment is provided under the open GPLv3 license: https://gitlab.com/hceye/HCeye/

Author Contributions

PT designed experiment, with assistance from NB, HS, and ZH. NB programmed experiment, with contributions and edits by PT. PT and ZH ran human subjects. ZH and PT programmed data analysis. PT wrote manuscript, with contributions from HS, ZH, and NB. HS supervised all research.

Funding

Supplementary Material

Detailed experimental procedures can be found online at http://journal.frontiersin.org/article/10.3389/fict.2015.00017

References

Allport, D., Styles, E., and Hsieh, S. (1994). "Shifting intentional set: exploring the dynamic control of tasks," in *Attention and Performance 15: Conscious and Non-conscious Information Processing*, eds C. Umilt and M. Moscovitch (Cambridge, MA: The MIT Press), 421–452.

Ashmore, M., Duchowski, A. T., and Shoemaker, G. (2005). "Efficient eye pointing with a fisheye lens," in *GI '05* (Waterloo, ON: Canadian Human-Computer Communications Society, School of Computer Science, University of Waterloo), 203–210.

Barber, D., Davis, L., Nicholson, D., Finkelstein, N., and Chen, J. Y. (2008). *The Mixed Initiative Experimental (MIX) Testbed for Human Robot Interactions with Varied Levels of Automation*. Technical Report, Orlando, FL: DTIC Document.

Bee, N., Prendinger, H., André, E., and Ishizuka, M. (2006). "Automatic preference detection by analyzing the gaze 'Cascade Effect'," in *Electronic Proceedings of the Second Conference on Communication by Gaze Interaction* (Turin: COGAIN), 63–66.

Bolt, R. A. (1981). "Gaze-orchestrated dynamic windows," in *SIGGRAPH '81* (New York, NY: ACM), 109–119.

Bosse, T., van Doesburg, W., van Maanen, P.-P., and Treur, J. (2007). "Augmented metacognition addressing dynamic allocation of tasks requiring visual attention," in *Foundations of Augmented Cognition*, eds D. Schmorrow and L. Reeves (Berlin: Springer), 166–175. doi:10.1007/978-3-540-73216-7_19

Bradshaw, J. L. (1968). Pupil size and problem solving. *Q. J. Exp. Psychol.* 20, 116–122. doi:10.1080/14640746808400139

Bulling, A., Roggen, D., and Troester, G. (2011). What's in the eyes for context-awareness? *IEEE Pervasive Comput.* 10, 48–57. doi:10.1109/MPRV.2010.49

Buscher, G., and Dengel, A. (2008). "Attention-based document classifier learning," in *Document Analysis Systems, 2008. DAS'08* (Nara: IEEE), 87–94.

Buscher, G., Dengel, A., and van Elst, L. (2008). "Query expansion using gaze-based feedback on the subdocument level," in *SIGIR '08* (New York, NY: ACM), 387–394. doi:10.1145/1390334.1390401

Buswell, G. (1935). *How People Look at Pictures: A Study of the Psychology and Perception in Art*. Oxford: Univ. Chicago Press.

Carlson, T., and Demiris, Y. (2009). "Using visual attention to evaluate collaborative control architectures for human robot interaction," in *Proceedings of New Frontiers in Human-Robot Interaction: A symposium at the AISB 2009 Convention* (Edinburgh: SSAISB), 38–43.

Castelhano, M. S., and Henderson, J. M. (2008). Stable individual differences across images in human saccadic eye movements. *Can. J. Exp. Psychol.* 62, 1–14. doi:10.1037/1196-1961.62.1.1

Core Team, R. (2013). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Couture-Beil, A., Vaughan, R. T., and Mori, G. (2010). "Selecting and commanding individual robots in a multi-robot system," in *Computer and Robot Vision (CRV), 2010 Canadian Conference* (Ottawa: IEEE), 159–166. doi:10.1109/CRV.2010.28

Coyne, J. T., Baldwin, C., Cole, A., Sibley, C., and Roberts, D. M. (2009). "Applying real time physiological measures of cognitive load to improve training," in *Foundations of Augmented Cognition. Neuroergonomics and Operational Neuroscience*, eds D. Schmorrow, I. Estabrooke, and M. Grootjen (Berlin: Springer), 469–478. doi:10.1007/978-3-642-02812-0_55

Craik, K. J. W. (1948). Theory of the human operator in control systems. *Br. J. Psychol.* 38, 142–148. doi:10.1111/j.2044-8295.1948.tb01149.x

De Greef, T., van Dongen, K., Grootjen, M., and Lindenberg, J. (2007). "Augmenting cognition: reviewing the symbiotic relation between man and machine," in *Foundations of Augmented Cognition*, eds D. Schmorrow and L. Reeves (Berlin: Springer), 439–448. doi:10.1007/978-3-540-73216-7_51

Debecker, J., and Desmedt, J. E. (1970). Maximum capacity for sequential one-bit auditory decisions. *J. Exp. Psychol.* 83, 366. doi:10.1037/h0028848

DeJong, B. P., Colgate, J. E., and Peshkin, M. A. (2011). "Mental transformations in human-robot interaction," in *Mixed Reality and Human-Robot Interaction*, ed. X. Wang (Netherlands: Springer), 35–51. doi:10.1007/978-94-007-0582-1_3

Elkind, J., and Sprague, L. (1961). Transmission of information in simple manual control systems. *IRE Trans. Hum. Fact. Electron.* HFE-2, 58–60. doi:10.1109/THFE2.1961.4503299

Feise, R. J. (2002). Do multiple outcome measures require p-value adjustment? *Med. Res. Methodol.* 2, 8. doi:10.1186/1471-2288-2-8

Fitts, P., and Posner, M. (1967). *Human Performance*. Oxford: Brooks/Cole.

Fono, D., and Vertegaal, R. (2004). "EyeWindows: using eye-controlled zooming windows for focus selection," in *Proceedings of UIST 2004. In Video Proceedings* (Santa Fe: ACM).

Fono, D., and Vertegaal, R. (2005). "EyeWindows: evaluation of eye-controlled zooming windows for focus selection," in *CHI '05* (New York, NY: ACM), 151–160. doi:10.1145/1054972.1054994

Funk, K., and McCoy, B., (1996). A functional model of flightdeck agenda management. *Proc. Hum. Fact. Ergon. Soc. Annu. Meet.* 40, 254–258. doi:10.1177/154193129604000424

Gelman, A., Hill, J., and Yajima, M. (2012). Why we (usually) don't have to worry about multiple comparisons. *J. Res. Educ. Effect.* 5, 189–211. doi:10.1080/19345747.2011.618213

Goldberg, J. H., and Schryver, J. C. (1993). Eye-gaze control of the computer interface: discrimination of zoom intent. *Proc. Hum. Fact. Ergon. Soc. Annu. Meet.* 37, 1370–1374. doi:10.1518/107118193784162272

Goldberg, J. H., and Schryver, J. C. (1995). Eye-gaze-contingent control of the computer interface: methodology and example for zoom detection. *Behav. Res. Methods Instrum. Comput.* 27, 338–350. doi:10.3758/BF03200428

Goldwater, B. (1972). Psychological significance of pupillary movements. *Psychol. Bull.* 77, 340–355. doi:10.1037/h0032456

Gopher, D. (1993). *Attention and Performance*. Cambridge, MA: MIT Press, 299–322.

Granholm, E., Asarnow, R., Sarkin, A., and Dykes, K. (1996). Pupillary responses index cognitive resource limitations. *Psychophysiology* 33, 457–461. doi:10.1111/j.1469-8986.1996.tb01071.x

Granholm, E., and Steinhauer, S. R. (2004). Pupillometric measures of cognitive and emotional processes. *Int. J. Psychophysiol.* 52, 1–6. doi:10.1016/j.ijpsycho.2003.12.001

Haapalainen, E., Kim, S., Forlizzi, J. F., and Dey, A. K. (2010). "Psycho-physiological measures for assessing cognitive load," in *Proceedings of the 12th ACM International Conference on Ubiquitous Computing, UbiComp '10* (New York, NY: ACM), 301–310. doi:10.1145/1864349.1864395

Hammer, J. (1999). "Human factors of functionality and intelligent avionics," in *Handbook of Human Factors in Aviation*, eds D. Garland, J. Wise, and V. Hopkin (Mahwah, NJ: Erlbaum), 549–565.

Hess, E., and Polt, J. (1964). Pupil size in relation to mental activity during simple problem-solving. *Science* 140, 1191.

Hild, J., Muller, E., Klaus, E., Peinsipp-Byma, E., and Beyerer, J. (2013). "Evaluating Multi-Modal Eye Gaze Interaction for Moving Object Selection," in *Proc. ACHI* (Nice: IARIA), 454–459.

Hwang, B., Jang, Y.-M., Mallipeddi, R., and Lee, M. (2013). Probing of human implicit intent based on eye movement and pupillary analysis for augmented cognition. *Int. J. Imaging Syst. Technol.* 23, 114–126. doi:10.1002/ima.22046

Hyona, J., Tommola, J., and Alaja, A.-M. (1995). Pupil dilation as a measure of processing load in simultaneous interpretation and other language tasks. *Q. J. Exp. Psychol. A* 48, 598–612. doi:10.1080/14640749508401407

Hyrskykari, A. (2006). *Eyes in Attentive Interfaces: Experiences from Creating iDict, a Gaze-Aware Reading Aid*. Ph.D. thesis, University of Tampere, Finland.

Hyrskykari, A., Majaranta, P., and Raiha, K.-J. (2003). "Proactive response to eye movements," in *INTERACT*, Vol. 3 (Amsterdam: IOS Press), 129–136.

Iqbal, S. T., and Bailey, B. P. (2004). "Using eye gaze patterns to identify user tasks," in *The Grace Hopper Celebration of Women in Computing*, 5–10.

Jacob, R. J. K. (1990). "What you look at is what you get: Eye movement-based interaction techniques," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems CHI '90* (New York: ACM), 11–18. doi:10.1145/97243.97246

Jacob, R. J. (1993a). Eye movement-based human-computer interaction techniques: toward non-command interfaces. *Adv. Hum. Comput. Interact.* 4, 151–190.

Jacob, R. J. (1993b). Hot topics-eye-gaze computer interfaces: what you look at is what you get. *Computer* 26, 65–66. doi:10.1109/MC.1993.274943

Jacob, R. J. K. (1991). The use of eye movements in human-computer interaction techniques: what you look at is what you get. *ACM Trans. Inf. Syst.* 9, 152–169. doi:10.1145/123078.128728

Jakob, R. (1998). "The use of eye movements in human-computer interaction techniques: what you look at is what you get," in *Readings in Intelligent User Interfaces*, eds M. T. Maybury, and W. Wahlster (Burlington: Morgan Kaufmann), 65–83.

Jie, L., and Clark, J. J. (2008). Video game design using an eye-movement-dependent model of visual attention. *ACM Trans. Multimedia Comput. Commun. Appl.* 4, 1–22. doi:10.1145/1386109.1386115

Johnson, M., Kulkarni, K., Raj, A., Carff, R., and Bradshaw, J. M. (2005). "Ami: an adaptive multi-agent framework for augmented cognition," in *Proceedings of the 11th International Conference in Human Computer Interaction, AUGCOG Conference*, 22–27.

Just, M. A., and Carpenter, P. A. (1976). Eye fixations and cognitive processes. *Cogn. Psychol.* 8, 441–480. doi:10.1016/0010-0285(76)90015-3

Just, M. A., Carpenter, P. A., and Miyake, A. (2003). Neuroindices of cognitive workload: neuroimaging, pupillometric and event-related potential studies of brain work. *Theor. Issues Ergon. Sci.* 4, 56–88. doi:10.1080/14639220210159735

Kahneman, D. (1973). *Attention and Effort*. Prentice-Hall.

Kahneman, D., and Beatty, J. (1966). Pupil diameter and load on memory. *Science* 154, 1584. doi:10.1126/science.154.3756.1583

Kahneman, D., and Beatty, J. (1967). Pupillary responses in a pitch-discrimination task. *Percept. Psychophys.* 2, 101. doi:10.3758/s13414-011-0263-y

Kumar, M., Winograd, T., and Paepcke, A. (2007). "Gaze-enhanced scrolling techniques," in *CHI EA '07* (New York, NY: ACM), 2531–2536. doi:10.1145/1240866.1241036

Laeng, B., Sirois, S., and Gredeback, G. (2012). Pupillometry: a window to the preconscious? *Perspect. Psychol. Sci.* 7, 18–27. doi:10.1177/1745691611427305

Laqua, S., Bandara, S. U., and Sasse, M. A. (2007). "GazeSpace: eye gaze controlled content spaces," in *BCS-HCI '07* (Swinton: British Computer Society), 55–58.

Liu, S. S., Rawicz, A., Rezaei, S., Ma, T., Zhang, C., Lin, K., et al. (2012). An eye-gaze tracking and human computer interface system for people with ALS and other locked-in diseases. *J. Med. Biol. Eng.* 32, 111–116. doi:10.5405/jmbe.813

Marshall, S. P. (2002). "The index of cognitive activity: measuring cognitive workload, in Human factors and power plants, 2002," in *Proceedings of the 2002 IEEE 7th Conference on (IEEE)*, 7–5.

Marshall, S. P., Pleydell-Pearce, C. W., and Dickson, B. T. (2003). "Integrating psychophysiological measures of cognitive workload and eye movements to detect strategy shifts, in system sciences, 2003," in *Proceedings of the 36th Annual Hawaii International Conference on (IEEE)*, 6.

Mathan, S., Dorneich, M., and Whitlow, S. (2005). "Automation etiquette in the augmented cognition context," in *Proceedings of the 11th International Conference on Human-Computer Interaction (1st AnnualAugmented Cognition International)* (Mahwah, NJ: Lawrence Erlbaum).

McConkie, G. W., and Rayner, K. (1975). The span of the effective stimulus during a fixation in reading. *Percept. Psychophys.* 17, 578–586. doi:10.3758/BF03203972

Megaw, E. D., and Richardson, J. (1979). Target uncertainty and visual scanning strategies. *Hum. Factors* 21, 303–315. doi:10.1177/001872087902100305

Menneer, T., Stroud, M. J., Cave, K. R., Li, X., Godwin, H. J., Liversedge, S. P., et al. (2012). Search for two categories of target produces fewer fixations to target-color items. *J. Exp. Psychol. Appl.* 18, 404–418. doi:10.1037/a0031032

Merten, C., and Conati, C. (2006). *Eye-Tracking to Model and Adapt to User Meta-Cognition in Intelligent Learning Environments*. ACM, 39–46.

Miellet, S., O'Donnell, P. J., and Sereno, S. C. (2009). Parafoveal magnification: visual acuity does not modulate the perceptual span in reading. *Psychol. Sci.* 20, 721–728. doi:10.1111/j.1467-9280.2009.02364.x

Miellet, S., Zhou, X., He, L., Rodger, H., and Caldara, R. (2010). Investigating cultural diversity for extrafoveal information use in visual scenes. *J. Vis.* 10, 21. doi:10.1167/10.6.21

Monajjemi, V. M., Wawerla, J., Vaughan, R. T., and Mori, G. (2013). "Hri in the sky: creating and commanding teams of uavs with a vision-mediated gestural interface," in *Intelligent Robots and Systems (IROS), 2013 IEEE/RSJ International Conference* (Tokyo: IEEE), 617–623.

Myers, J. L., Well, A. D., and Jr, R. F. L. (2010). *Research Design and Statistical Analysis*, Third Edn. New York, NY: Routledge.

Nielsen, J., and Pernice, K. (2010). *Eyetracking Web Usability*. Berkeley, CA: New Riders.

Ohno, T. (2004). "EyePrint: support of document browsing with eye gaze trace," in *ICMI '04* (New York, NY: ACM), 16–23. doi:10.1145/1027933.1027937

Ohno, T. (2007). EyePrint: using passive eye trace from reading to enhance document access and comprehension. *Int. J. Hum. Comput. Interact.* 23, 71–94. doi:10.1080/10447310701362934

Parasuraman, R. (1986). "Vigilance, monitoring, and search," in *Handbook of Perception and Human Performance, Vol. 2: Cognitive Processes and Performance*, eds K. R. Boff, L. Kaufman, and J. P. Thomas (Oxford: John Wiley & Sons), 1–39.

Pavel, M., Wang, G., and Li, K. (2003). "Augmented cognition: allocation of attention, in system sciences, 2003," in *Proceedings of the 36th Annual Hawaii International Conference on (IEEE)* (Washington, DC: IEEE), 6.

Perneger, T. V. (1998). What's wrong with Bonferroni adjustments. *BMJ* 316, 1236–1238. doi:10.1136/bmj.316.7139.1236

Piquado, T., Isaacowitz, D., and Wingfield, A. (2010). Pupillometry as a measure of cognitive effort in younger and older adults. *Psychophysiology* 47, 560–569. doi:10.1111/j.1469-8986.2009.00947.x

Qvarfordt, P., and Zhai, S. (2005). "Conversing with the user based on eye-gaze patterns," in *CHI '05* (New York, NY: ACM), 221–230. doi:10.1145/1054972. 1055004

Raley, C., Stripling, R., Kruse, A., Schmorrow, D., and Patrey, J. (2004). Augmented cognition overview: improving information intake under stress. *Proc. Hum. Fact. Ergon. Soc. Annu. Meet.* 48, 1150–1154. doi:10.1177/154193120404801001

Rayner, K., and Bertera, J. H. (1979). Reading without a fovea. *Science* 206, 468–469. doi:10.1126/science.504987

Reder, S. M. (1973). On-line monitoring of eye-position signals in contingent and noncontingent paradigms. *Behav. Res. Methods Instrument.* 5, 218–228. doi:10. 3758/BF03200168

Rosch, J. L., and Vogel-Walcutt, J. J. (2013). A review of eye-tracking applications as tools for training. *Cogn. Technol. Work* 15, 313–327. doi:10.1007/s10111-012-0234-7

Rothman, K. J. (1990). No adjustments are needed for multiple comparisons. *Epidemiology* 1, 43–46. doi:10.1097/00001648-199001000-00010

Roy, D., Ghitza, Y., Bartelma, J., and Kehoe, C. (2004). "Visual memory augmentation: using eye gaze as an attention filter," in *Eighth International Symposium on ISWC 2004* Vol. 1. Washington, DC: IEEE, 128–131.

Russo, M. B., Stetz, M. C., and Thomas, M. L. (2005). Monitoring and predicting cognitive state and performance via physiological correlates of neuronal signals. *Aviat. Space Environ. Med.* 76(Suppl. 1), C59–C63.

Salvucci, D. D. (1999). "Inferring intent in eye-based interfaces: tracing eye movements with process models," in *CHI '99* (New York, NY: ACM), 254–261. doi:10. 1145/302979.303055

Sanbonmatsu, D. M., Strayer, D. L., Medeiros-Ward, N., and Watson, J. M. (2013). Who multi-tasks and why? Multi-tasking ability, perceived multi-tasking ability, impulsivity, and sensation seeking. *PLoS ONE* 8:e54402. doi:10.1371/journal. pone.0054402

Saville, D. J. (1990). Multiple comparison procedures: the practical solution. *Am. Stat.* 44, 174–180. doi:10.2307/2684163

Shimojo, S., Simion, C., Shimojo, E., and Scheier, C. (2003). Gaze bias both reflects and influences preference. *Nat. Neurosci.* 6, 1317–1322. doi:10.1038/nn1150

Sibert, L. E., and Jacob, R. J. (2000). "Evaluation of eye gaze interaction," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '00* (New York: ACM), 281–288. doi:10.1145/332040.332445

Simpson, H., and Hale, S. (1969). Pupillary changes during a decision-making task. *Percept. Mot. Skills* 29, 495–498. doi:10.2466/pms.1969.29.2.495

Spakov, O., and Miniotas, D. (2005). "Gaze-based selection of standard-size menu items," in *ICMI '05* (New York, NY: ACM), 124–128. doi:10.1145/1088463. 1088486

St. John, M., Kobus, D., Morrison, J., and Schmorrow, D. (2004). Overview of the DARPA augmented cognition technical integration experiment. *Int. J. Hum. Comput. Interact.* 17, 131–149. doi:10.1207/s15327590ijhc1702_2

St John, M., Kobus, D. A., and Morrison, J. G. (2003). *DARPA Augmented Cognition Technical Integration Experiment (TIE)*. Technical Report, DTIC Document.

Starker, I., and Bolt, R. A. (1990). "A gaze-responsive self-disclosing display," in *CHI '90* (New York, NY: ACM), 3–10. doi:10.1145/97243.97245

Sundstedt, V. (2012). Gazing at games: an introduction to eye tracking control. *Synth. Lect. Comput. Graph. Anim.* 5, 1–113. doi:10.2200/S00395ED1V01 Y201111CGR014

Tanriverdi, V., and Jacob, R. J. K. (2000). "Interacting with eye movements in virtual environments," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, CHI '00* (New York: ACM), 265–272. doi:10.1145/332040. 332443

Taylor, P., He, Z., Bilgrien, N., and Siegelmann, H. T. (2015). Human strategies for multitasking, search, and control improved via real-time memory aid for gaze location. *Hum. Media Interact.* 2:15. doi:10.3389/fict.2015.00015

Taylor, R. M., Brown, L., and Dickson, B. (2003). *From Safety Net to Augmented Cognition: Using Flexible Autonomy Levels for On-line Cognitive Assistance and Automation*. Technical Report, DTIC Document.

Taylor, T., Pradhan, A., Divekar, G., Romoser, M., Muttart, J., Gomez, R., et al. (2013). The view from the road: the contribution of on-road glance-monitoring technologies to understanding driver behavior. *Accid. Anal. Prev.* 58, 175–186. doi:10.1016/j.aap.2013.02.008

Treisman, A., and Gelade, G. (1980). A feature-integration theory of attention. *Cogn. Psychol.* 12, 97–136. doi:10.1016/0010-0285(80)90005-5

Van Gerven, P. W. M., Paas, F., Van Merrienboer, J. J. G., and Schmidt, H. G. (2004). Memory load and the cognitive pupillary response in aging. *Psychophysiology* 41, 167–174. doi:10.1111/j.1469-8986.2003.00148.x

Ververs, P. M., Whitlow, S., Dorneich, M., and Mathan, S. (2005). "Building Honeywell's adaptive system for the augmented cognition program," in *1st International Conference on Augmented Cognition* (Las Vegas, NV).

Wankhede, H. S., Chhabria, S. A., and Dharaskar, R. V. (2013). Human computer interaction using eye and speech: the hybrid approach. *Int. J. Emer. Sci. Eng.* 1, 54–58.

Watson, J. M., and Strayer, D. L. (2010). Supertaskers: profiles in extraordinary multitasking ability. *Psychon. Bull. Rev.* 17, 479–485. doi:10.3758/PBR.17.4.479

Wickens, C. D. (1997). *Computational Models of Human Performance in the Design and Layout of Controls and Displays*. Wright-Patterson AFB: Crew System Ergonomics Information Analysis Center (CSERIAC).

Wiener, E. L., and Curry, R. E. (1980). Flight-Deck automation: promises and problems. *Ergonomics* 23, 995–1011. doi:10.1080/00140138008924809

Wierda, S. M., Rijn, H. V., Taatgen, N. A., and Martens, S. (2012). Pupil dilation deconvolution reveals the dynamics of attention at high temporal resolution. *Proc. Natl. Acad. Sci. U.S.A.* 109, 8456–8460. doi:10.1073/pnas.1201858109

Xu, S., Jiang, H., and Lau, F. C. (2008). "Personalized online document, image and video recommendation via commodity eye-tracking," in *RecSys '08* (New York, NY: ACM), 83–90. doi:10.1145/1454008.1454023

Yarbus, A. L. (1967). *Eye Movements and Vision*. New York, NY: Plenum Press.

Zekveld, A. A., Heslenfeld, D. J., Johnsrude, I. S., Versfeld, N. J., and Kramer, S. E. (2014). The eye as a window to the listening brain: neural correlates of pupil size as a measure of cognitive listening load. *Neuroimage* 101, 76–86. doi:10.1016/j.neuroimage.2014.06.069

Zekveld, A. A., and Kramer, S. E. (2014). Cognitive processing load across a wide range of listening conditions: insights from pupillometry: processing load across a wide range of listening conditions. *Psychophysiology* 51, 277–284. doi:10.1111/ psyp.12151

Zhai, S., Morimoto, C., and Ihde, S. (1999). "Manual and gaze input cascaded (MAGIC) pointing," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (ACM), 246–253.