How Coordinated Visual Attention on a Target Area of Interest is Impacted by a Change in Workload Over Time

Shannon P. Devlin[1], Jake R. Flynn[2], and Sara L. Riggs[1]
University of Virginia[1], Clemson University[2]

Data-rich environments rely on operators to collaborate, especially in light of workload changes. This work explores the relationship between the operators' shared visual attention patterns on a target area of interest (AOI), i.e. the AOI causing a workload change, and how it affects collaborative performance. Eye tracking data was collected from ten pairs of participants who completed two scenarios, the first being low workload and the second being high workload, in an unmanned aerial vehicle (UAV) command and control testbed. Then, best and worst performing pairs were compared in terms of two shared visual attention metrics: (1) percent gaze overlap and (2) the phi coefficient for the target AOI. The results showed that coordinated visits to and from the target AOI were associated with better performance during high workload. These results suggest including quantitative measures of visual attention can be indicators of the adaptation process in real-time.

INTRODUCTION

Dynamic and data-rich domains rely on multiple operators to coordinate and complete tasks. Operators may struggle to manage these tasks, especially when there are changes in cognitive workload. For example, accidents in unmanned aerial vehicle (UAV) command and control have occurred when operators were not fully collaborating during transitions between low and high workload (Williams, 2006). Quantitative measures can potentially be used to assess collaboration in real-time and prevent accidents. Of which, eye tracking technology is a promising means to do so given: (a) the majority of the information in these types of domains include various visual displays and (b) studying visual attention patterns can provide insights on changes in cognitive load (Coral, 2016) and task completion processes (e.g., Chierichetti, Kumar, Raghavan, & Sarlos, 2012). However, research is needed to understand whether eye tracking can capture aspects related to successful collaboration, i.e., how their shared visual attention patterns change when there are changes in the environment.

Studying the visual attention of multiple people collaborating in a shared environment has been examined in various contexts (Böckler, Knoblich, Sebanz, & Böckler, 2012; Brennan, Chen, Dickinson, Neider, & Zelinsky, 2008; D'Angelo & Begel, 2017). Research has shown that visual attention is influenced by the presence of another individual, but this influence depends on the context. For example, one commonly used metric to characterize shared visual attention is *percent gaze overlap,* which quantifies the amount of time multiple observers are concurrently viewing the same predetermined area of the display (Pietinen, Bednarik, & Tukiainen, 2010). Increases in percent gaze overlap have corresponded with improved performance (Gergle & Clark, 2011; Hajari, Cheng, Zheng, & Basu, 2016), but this is not always the case (e.g. Villamor & Rodrigo, 2018).

Other measures may provide more insights about collaboration. For example, in lag sequential analysis, two time series in a dynamic system are studied quantifiably at different lag times to understand the relationship between the two time series (Bakeman & Gottman, 1997). A metric associated with lag sequential analysis is the phi coefficient (ϕ), which quantifies how much coordination there is between two time series as a function of visits to and from a specific *state* in the system. When studying two people's visual attention as they collaborate, the two time series in the system consists of the scanpaths generated by each person. In this work, scanpaths are based on the order in which *areas of interest* (AOIs) on the display are visited, making AOIs the states of the system. Therefore, the phi coefficient can provide a real-time quantitative measure of how two people are coordinating their scanpaths as a function of visits to and from a *target AOI,* i.e. a predetermined AOI on the display.

The phi coefficient is calculated as a function of when the scanpaths are and are not at a target AOI. The phi coefficient is typically calculated for different lag times to see if there is leader/follower behavior between collaborators (Coco & Dale, 2014). Table 1 provides an example of two participants' scanpaths (i.e. time series) over a 5 s window, assuming a sampling rate of 1 s, on a display that has four AOIs. It indicates when the target AOI—AOI 3 in this example—is matched or mismatched between the two scanpaths for when there is no lag and a 1 second lag, (i.e. a shift of participant 2's scanpath by one time period to the right).

TABLE 1: Hypothetical scanpath of two participants viewing a display with four AOIs for a 5 s time window and an indication of when scanpaths match at the target AOI (AOI 3)

	5 s Time Window				
	1s	*2s*	*3s*	*4s*	*5s*
Participant 1	1	4	3	3	3
Participant 2	2	3	3	1	3
Participant 2' (1 s lag)		2	3	3	1
Participant 1 & 2 match w.r.t target AOI?	Yes, both **not** at AOI 3	No	Yes, both at AOI 3	No	Yes, both at AOI 3
Participant 1 & 2' match w.r.t target AOI?	N/A	Yes, both **not** at AOI 3	Yes, both at AOI 3	Yes, both at AOI 3	No

Calculating the phi coefficient requires creating a contingency table that shows the frequency of when the two scanpaths are matching and mismatching with respect to the target AOI. Given the phi coefficient is calculated based on visits to one target AOI, the contingency table only has two rows and two columns: one representing the target AOI and the other being a consolidation of all other AOIs. The rows will be for one participant's scanpath and the columns will be for the other participant's scanpath (Table 2). The table would then be populated with number of instances when both scanpaths are at the target AOI (*A* in Table 2), when both are not (*D* in Table 2), and when one is and one is not (*B* and *C,* respectively, in Table 2).

TABLE 2: Contingency table used to calculate phi coefficient

Scanpath 2 \ Scanpath 1	Target AOI	All other AOIs
Target AOI	*A*	*B*
All other AOIs	*C*	*D*

Subsequently, the general formula of the phi coefficient, where A, B, C, and D are the counts from the contingency table (Table 2), is:

$$\phi = \frac{AD - BC}{\sqrt{(A+B)(C+D)(A+C)(B+D)}} \quad (1)$$

Values for the phi coefficient range from -1 to +1, where -1 indicates a perfect negative association between visits to the target AOI and overall coordinated attention (i.e. visits to and from the target AOI are detracting from coordination, all counts are from B and/or C) and +1 being a perfect positive association between the target AOI and overall coordinated attention (i.e. visits to and from the AOI are directly increasing coordinated visual attention, all counts are from A and/or D). The strength of this association are determined by the phi coefficient's absolute values and are interpreted as follows: 0-0.05 as none or very weak, 0.05-0.10 as weak, 0.10-0.15 as moderate, 0.15-0.25 as strong, and any magnitude above 0.25 as very strong (Akoglu, 2018). For the example from Table 1, the phi coefficient for participants 1 and 2 would be 0.17 (A=2, B=1, C=1, D=1) and for participants 1 and 2' is 0.58 (A=2, B=0, C=1, D=1). This suggests that there is a *strong* positive association between coordinated visits to and from AOI 3 and overall coordination for participants 1 and 2 and a *very strong* positive association for participants 1 and 2 when participant 2's scanpath is lagged 1 s. This example shows that the phi coefficient can capture nuanced coordination patterns between two collaborators, such as coordinating views to the target AOI not being perfectly synchronized (e.g., participants 1 and 2') – a nuance percent gaze overlap cannot capture.

Previous research examining shared visual attention has focused on performance; however, the effect a workload change has on shared visual attention strategies and its subsequent impact on performance has received less attention. For example, studying the team adaptation process, i.e. actions in response to an environmental change (Maynard & Kennedy, 2016), has been typically completed post-hoc and is qualitative in nature. Quantitative measures, such as shared visual attention, could shed new light on the adaption process (Resick et al., 2010). These measures could also be studied more granularly, therefore leveraging the design of technology to support real-time collaboration (Fiore & Wiltshire, 2016; Fussell, Kraut, & Siegel, 2000).

This research builds on previous work that examined whether shared visual attention affected the collaboration strategies of successful and unsuccessful pairs managing a workload change (Devlin, Flynn, & Riggs, 2019). Our previous work found the most successful pairs of participants adapted both their task completion and shared visual attention strategies more readily upon experiencing a change in workload. Specifically, they substantially increased their shared visual attention on the AOI causing the workload change. This present work further examines how coordinated visual attention to and from this AOI impacts performance by measuring percent gaze overlap and the phi coefficient over time for both best and worst performing pairs. The goal here is to quantify the adaptation process of pairs over time. This in turn could be used to inform technology design in the future (Fiore & Wiltshire, 2016). The chosen application for this work is UAV command and control given their aim to incorporate quantitative measures in their technology design to facilitate effective collaboration in real-time (Sibley, Coyne, & Morrison, 2015).

METHOD

Participants

Ten pairs of undergraduate students (20 students total) at Clemson University were recruited for the study (*M* = 21.3 years, *SE* = .24 years). Each pair consisted of one male and one female who did not previously know each other. The experiment lasted from 75-90 minutes and participants were compensated $10/hour for their time.

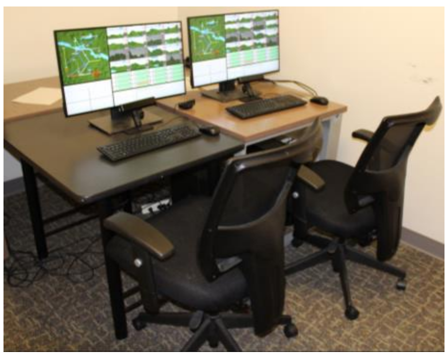
Experimental Setup

The design of the simulation was based on the 'Vigilant Spirit Control Station' the U.S. Air Force uses to develop interfaces to control multiple UAVs (Feitshans, Rowe, Davis, Holland, & Berger, 2008). The simulation was developed using the Unity gaming engine and ran on a desktop computer (28" monitor, 2560×1440 screen resolution). Pairs were collocated, but each participant viewed separate monitors and used separate mice to input responses (Figure 1a). The simulation was networked so participants could see inputs from their partner in real-time (e.g., when participant 1 responded to a chat message, participant 2 could see his/her response in real-time).

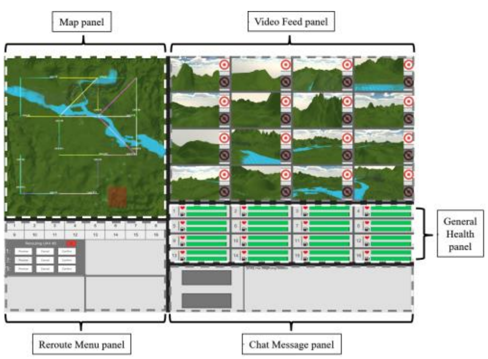
Two desktop mounted FOVIO eye trackers with a sampling rate of 60 Hz were used to collect point of gaze data. One eye tracker was placed below each monitor and participants sat 26-28 inches from the monitor. The average degree of error for this eye tracker is 0.78° (SD = 0.59°; Eyetracking, 2011).

Tasks

Each pair was responsible for completing one primary task and three secondary tasks for up to 16 UAVs (Figure 1b). Although all tasks were the pair's responsibility, only one participant from each pair had to complete each task. The primary task was the target detection task where pairs monitored each UAV's video feed and indicated whether a target (i.e., a semi-transparent cube) was present. This task took place in the video feed panel. The secondary tasks included a rerouting task (avoiding no-fly zones), fuel leak task (maintaining UAV health), and chat message task (responding to chat messages). More details can be found in (Devlin et al., 2019).



(a)



(b)

Figure 1. (a) The experimental setup with two networked computers simultaneously running the experimental testbed (b) Screenshot of the experimental testbed with panels labelled

Workload

Workload was manipulated by varying the number of UAVs the pair of participants needed to simultaneously manage for the primary task (i.e., target detection task). There were two workload conditions: low and high. For the low workload condition, the pair was responsible for 3-5 UAVs at all times and for the high workload condition they were responsible for 13-16 UAVs at all times. Pairs always completed the low workload condition before the high workload condition. Due to workload being manipulated via the rate of the primary task, the video feed panel is designated as the target AOI for both shared visual attention metrics.

Procedure

Participants of each pair read and signed the consent form and were then briefed about the study's goals and task expectations. Participants then independently completed a five-minute training session. By the end of the training session, participants had to demonstrate they could achieve 70% accuracy for all tasks. The pairs were then informed on how the simulation was networked and were then provided three minutes to introduce themselves to one another and discuss anything they deemed necessary. The pairs then completed the low workload condition, were provided a short break, and then completed the high workload condition. Each condition was a 15-minute testbed scenario.

Experimental Design

The independent variable in this study was pair performance (best performing vs. worst performing). Best and worst performing pairs were determined based on the total points scored in the low and high workload scenario. The best performing pairs were the three highest scoring pairs and the worst performing pairs were the three lowest scoring pairs. More details can be found in (Devlin et al., 2019). Dependent variables included the two shared visual attention metrics of the target AOI (i.e., video feed panel; Figure 1b): percent gaze overlap and the mean of the maximum absolute value of the phi coefficient for each minute of both scenarios.

RESULTS

The gaze data was screened to meet data quality requirements as outlined in ISO/TS 15007-2:2014-09, which states that at most 15% data loss is acceptable. Following this guideline, no participants were excluded from the study and the mean data loss was 9.23%. Given, previous work showed that the best performing pairs had higher percent gaze overlap on the target AOI (Devlin et al., 2019), percent gaze overlap was then calculated for each minute of the scenario for the target AOI only. The best performing pairs' percent gaze overlap ranged for each scenario (low workload range: 17.4-33.0%; high workload range: 40.0-63.7%) and were consistently higher than worst performing pairs, (low workload range: 1.0-7.3%; high workload range: 9.9-22.2%). These ranges show best performing pairs not only had higher percent gaze overlap overall, but they also increased these levels more from low to high workload.

The phi coefficient for the target AOI (i.e., video feed panel) was calculated for each minute in both scenarios to understand how visits to and from the target AOI impacted shared visual attention patterns and how this impact evolves over time during a workload change. The largest positive or negative phi coefficient value was of interest as this indicated when the pair's visual attention coordination was most impacted by the target AOI. Both values were found for each minute by calculating the phi coefficient for each lag in a ±10 s window (recommended by Dale, Kirkham, et al., 2011) across both best and worst performing pairs' scanpaths. Then the mean for both the largest positive and negative values was calculated, with the maximum absolute value being plotted for each minute for both scenarios (Figure 2 and 3). For the majority of the low workload scenario, both the best and worst performing pairs' largest phi coefficient was negative (13 or 14 minutes of the 15-minute scenario, respectively) and had either moderate, strong, or very strong associations (Akoglu, 2018; Figure 2). The worst performing pairs' largest phi coefficient remained relatively consistent during the low workload scenario (except for the last minute of low workload), whereas the best performing pairs' largest values fluctuated more over time.

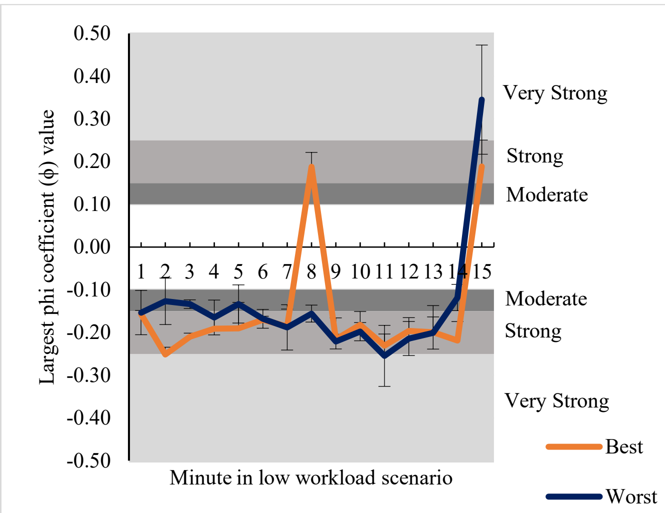


Figure 2. Mean maximum phi coefficient magnitude for video feed panel over time for best and worst performing pairs during the low workload scenario

For the high workload scenario, the best performing pairs had more positive phi coefficient values (10 minutes of the 15-minute scenario) whereas the worst performing pairs did not (5 minutes of the 15-minute scenario; Figure 3). All associations were moderate, strong, or very strong. This suggests the coordination of best performing pairs was a function of coordinated views to and from the target AOI during high workload.

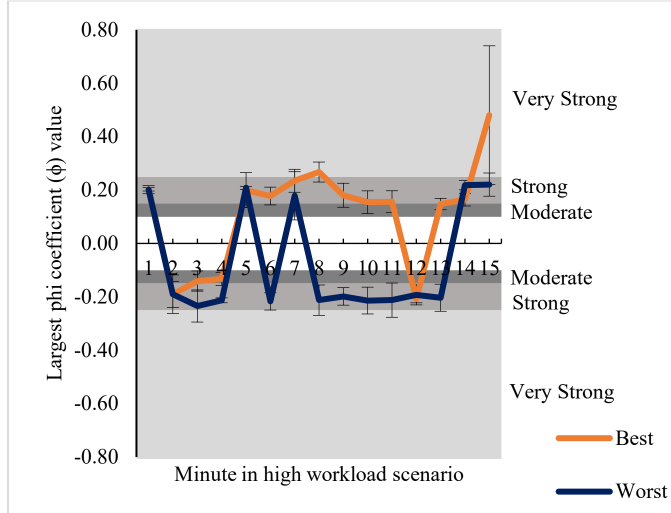


Figure 3. Mean maximum phi coefficient magnitude for the video feed panel over time for best and worst performing pairs during the high workload scenario

For each phi coefficient value, there was also an associated lag time: *negative* lag times correspond with leading coordinated behavior whereas *positive* lag times corresponding with lagging coordinated behavior. For the analysis here, a lag time could change sign a minimum of zero times (i.e., be the same sign the entire scenario) or a maximum of 14 times (i.e., switch every minute of the scenario). We found no trend in these lag times as the sign of the lag time changed frequently over the course of each scenario for the best and worst performing pairs. During the low workload scenario, the sign of the lag time changed, on average, 7 times for best performing pairs and 8 times for worst performing pairs. For high workload, the sign changed, on average, 6 times for best performing pairs and 5 times for worst performing pairs. The lag time values were also less or greater than zero, suggesting there was always either leading or lagging behavior.

DISCUSSION & CONCLUSION

This work aimed to understand how shared visual attention of pairs changed over time when workload changed from low to high. The results show there is a positive relationship between gaze overlap and performance over time, supporting previous work (Devlin et al., 2019; Gergle & Clark, 2011; Hajari et al., 2016). It specifically suggests performance is improved when participants substantially increase their shared visual attention on the AOI causing a workload change, (i.e. the target AOI) and sustain those levels over time.

The results from the phi coefficient analysis revealed differences between best and worst performing pairs when workload changed. There were minimal differences between best and worst performing pairs during low workload as the majority of the largest phi coefficient values had at least a moderate negative association between visits to the target AOI and overall coordinated visual attention. This suggests strategically coordinating visual attention to and from the target AOI was not critical for coordinating attention overall or for improving performance during low workload. However, this was not the case for high workload. Here the best performing pairs largest phi coefficient values were positive, suggesting coordinated visual attention to and from the target AOI was associated with higher levels of coordination overall. This suggests improved performance during a workload increase is not only dependent on viewing the target AOI at the same time, (as indicated by percent gaze overlap), but also coordinating visual attention to and from it, (as indicated by phi coefficient). Given the worst performing pairs' largest phi coefficient values remained similar between low and high workload, our results show performance suffers when pairs do not adapt specifically and strategically to the workload increase (Maynard & Kennedy, 2016). This study provides a potentially promising quantitative measure that can assess the absence of adaptation in real-time, which is lacking in the adaptation process literature. We do want to note that the applicability of this analysis is context dependent on what is defined as the target AOI.

Our results showed there was no set pattern of the lag times associated with the largest phi coefficient values for the best and worst performing pairs. This finding may be attributed to the fact we did not assign roles to each participant, which is unlike previous studies. However, this type of collaboration is expected to be the structure of future UAV command and control (Sibley et al., 2015). Although more work is needed, incorporating lag times could help inform *how* a pair is coordinating their visual attention towards a target AOI and be used to inform *how* the technology should intervene to improve coordination. With our initial analysis of lag time values, we found pairs often switched between leading and lagging visual attention behavior on the target AOI. One possible explanation of this finding may be due to participants taking turns on who primarily attends to the target AOI when another AOI needs attention. Future work could further extend on this preliminary finding to better understand how these pairs' specific approach to managing a workload change impacted performance. For example, our initial analysis found worst performing pairs number of switches slightly decreased from low to high workload while best performing pairs' number of switches remained similar. This potentially suggests continual dynamic diversification of task responsibilities may be part of the observed performance advantage.

Overall, this body of work shows that when workload increases, coordinating shared visual attention improves performance. The findings support the potential of using eye tracking metrics such a percent gaze overlap and phi coefficient in real-time to inform and improve collaboration. Future research needs to explore how to effectively use and present this information (e.g., the impact of seeing a partner's gaze in real-time; Schneider et al., 2018). These findings also better inform what constitutes as a successful adaptation process by using a quantitative, real-time measure. This can help strengthen both the adaptation process literature and technology design.

REFERENCES

Akoglu, H. (2018). User's guide to correlation coefficients. *Turkish Journal of Emergency Medicine*, *18*(3), 91–93.

Bakeman, R., & Gottman, J. M. (1997). *Observing interaction: An introduction to sequential analysis*. Cambridge University Press.

Böckler, A., Knoblich, G., Sebanz, N., & Böckler, A. (2012). Effects of coactor's focus of attention on task performance. *Journal of Experimental Psychology: Human Perception and Performance*.

Brennan, S. E., Chen, X., Dickinson, C. A., Neider, M. B., & Zelinsky, G. J. (2008). Coordinating cognition: The costs and benefits of shared gaze during collaborative search. *Cognition*, *106*(3), 1465–1477.

Chierichetti, F., Kumar, R., Raghavan, P., & Sarlos, T. (2012). Are web users really Markovian? *Proceedings of the 21st International Conference on World Wide Web - WWW '12*, 609–618.

Coco, M. I., & Dale, R. (2014). Cross-recurrence quantification analysis of categorical and continuous time series: An R package. *Frontiers in Psychology*, *5*(510), 1–14.

Coral, M. P. (2016). *Analyzing cognitive workload through eye-related measurements: A meta-analysis*. Wright State University.

D'Angelo, S., & Begel, A. (2017). Improving communication between pair programmers using shared gaze awareness. *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems - CHI '17*, 6245–6290.

Dale, R., Kirkham, N. Z., & Richardson, D. C. (2011). The dynamics of reference and shared visual attention. *Frontiers in Psychology*, *2*, 1–11.

Devlin, S. P., Flynn, J. R., & Riggs, S. L. (2019). Examining the visual attention of pairs of operators during a low to high workload change. *Proceedings of the Human Factors and Ergonomics Society 2019 Annual Meeting*, *1*(2018), 201–205.

Feitshans, G., Rowe, A., Davis, J., Holland, M., & Berger, L. (2008). Vigilant spirit control station (vscs) the face of counter. *AIAA Guidance, Navigation and Control Conference and Exhibit*.

Fiore, S. M., & Wiltshire, T. J. (2016). Technology as teammate: Examining the role of external cognition in support of team cognitive processes. *Frontiers in Psychology*, *7*(1531), 1–17.

Fussell, S. R., Kraut, R. E., & Siegel, J. (2000). Coordination of communication: Effects of shared visual context on collaborative work. *CSCW '00*. Philadelphia, PA.

Gergle, D., & Clark, A. T. (2011). See what I'm saying? Using dyadic mobile eye tracking to study collaborative reference. *Proceedings of the ACM 2011 Conference on Computer Supported Cooperative Work - CSCW '11*, 435.

Hajari, N., Cheng, I., Zheng, B., & Basu, A. (2016). Determining team cognition from delay analysis using cross recurrence plot. *38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 3482–3485. Orlando, FL.

Maynard, M. T., & Kennedy, D. M. (2016). *Team adaptation and resilience: What do we know and what can be applied to long-duration isolated, confined, and extreme contexts*. Houston, TX.

Pietinen, S., Bednarik, R., & Tukiainen, M. (2010). Shared visual attention in collaborative programming: A descriptive analysis. *Proceedings of the 2010 ICSE Workshop on Cooperative and Human Aspects of Software Engineering - CHASE '10*, 21–24. Cape Town, South Africa.

Resick, C. J., Murase, T., Bedwell, W. L., Sanz, E., Jiménez, M., & Dechurch, L. A. (2010). *Mental model metrics and team adaptability: A multi-facet multi-method examination*. *14*(4), 332–349.

Schneider, B., Sharma, K., Cuendet, S., Zufferey, G., Dillenbourg, P., & Pea, R. (2018). Leveraging mobile eye-trackers to capture joint visual attention in co-located collaborative learning groups. *International Journal of Computer Supported Collaborative Learning*.

Sibley, C., Coyne, J., & Morrison, J. (2015). Research considerations for managing future unmanned systems. *AAAI Spring Symposium*, 54–59.

Villamor, M., & Rodrigo, M. M. (2018). Predicting successful collaboration in a pair programming eye tracking experiment. *Adjunct Publication of the 26th Conference on User Modeling, Adaptation and Personalization*, 263–268.

Williams, K. W. (2006). *Human factors implications of unmanned aircraft accidents: Flight-control problems*.