Author-accepted manuscript

Eye Tracking During A Complex Aviation Task For Insights Into Information Processing
N Ayala, S Kearns, E Irving, S Cao, E Niechwiej-Szwedo
Journal of Visualized Experiments (JoVE), e66886, 2025

TITLE:
Eye Tracking During A Complex Aviation Task For Insights Into Information Processing

AUTHORS AND AFFILIATIONS:
Naila Ayala[1,5], Suzanne Kearns[2,5], Elizabeth Irving[3,5], Shi Cao[4,5], Ewa Niechwiej-Szwedo[1,5]

[1]Department of Kinesiology, University of Waterloo, Waterloo, ON, Canada
[2]Department of Geography and Aviation, University of Waterloo, Waterloo, ON, Canada
[3]School of Optometry and Vision Science, University of Waterloo, Waterloo, ON, Canada
[4]Department of Systems Design Engineering, University of Waterloo, Waterloo, ON, Canada
[5]Waterloo Institute for Sustainable Aeronautics, University of Waterloo, Waterloo, ON, Canada

Email addresses of the co-authors:

Naila Ayala	naila.ayala@uwaterloo.ca
Suzanne Kearns	suzanne.kearns@uwaterloo.ca
Elizabeth Irving	elizabeth.irving@uwaterloo.ca
Shi Cao	shi.cao@uwaterloo.ca

Corresponding author:

Ewa Niechwiej-Szwedo	eniechwiej@uwaterloo.ca

SUMMARY:
Eye tracking is a non-invasive method to probe information processing. This article describes how eye tracking can be used to study gaze behavior and understand how underlying cognitive, perceptual, and sensorimotor processes support action and performance during a flight simulation task in low-time pilots (i.e., <350 flight hours).

ABSTRACT:
Eye tracking has been used extensively as a proxy to gain insight into the cognitive, perceptual, and sensorimotor processes that underlie skill performance. Previous work has shown that traditional and advanced gaze metrics reliably demonstrate robust differences in pilot expertise, cognitive load, fatigue, and even situation awareness (SA) primarily through highly controlled laboratory-based paradigms where head movements are restrained.

This manuscript describes the methodology and contributions of a wearable eye tracker and gaze mapping algorithm that captures naturalistic head and eye movements (i.e., gaze) in ecological environments (i.e., a high-fidelity flight simulator). The introduction of this method extends the wearable eye tracker capabilities in generating area of interest (AOI)-based gaze analyses that provide more context related to where participants are looking, why they are looking there, and how efficiently they are processing the fixated information. As such, the current study examined the utility of a wearable eye tracker and computer vision algorithm to assess changes in gaze behavior in response to manipulations in task difficulty (i.e., absence or presence of an in-flight emergency). In-flight emergency scenarios are scantily researched and provide a significantly challenging task manipulation to assess changes in information processing.

The results demonstrated that gaze was significantly impacted when the emergency event was introduced. Specifically, attention allocation, gaze dispersion, and gaze sequence complexity significantly decreased and became highly concentrated on looking outside the front window and at the airspeed gauge during the emergency scenario (all p values < 0.05). The utility and limitations of employing a wearable eye tracker in a high-fidelity flight simulation environment to understand the spatiotemporal characteristics of gaze behavior and its relation to information processing in the aviation domain are discussed.

INTRODUCTION:

Humans predominantly interact with the world around them by first moving their eyes and head to focus their line of sight (i.e., gaze) toward a specific object or location of interest. This is particularly true in complex environments such as aircraft cockpits where pilots are faced with multiple competing stimuli. Gaze movements enable the collection of high-resolution visual information that allows humans to interact with their environment in a safe and flexible manner[1], which is of paramount importance in aviation. Seminal work has demonstrated that gaze behavior is highly task-dependent, tightly linked to underlying perceptual, cognitive, and motor processes[1,2,3], and, thus, has a direct influence on the planning and execution of subsequent actions[3]. As such, the analysis of gaze behavior during aviation tasks provides a sensitive and non-invasive method to examine how specific eye movement patterns relate to various aspects of information processing and performance.

Several studies have demonstrated an association between gaze and task performance across various laboratory paradigms, as well as complex real-world tasks (i.e., operating an aircraft). For instance, task-relevant areas tend to be fixated more frequently and for longer total durations, suggesting that fixation location, frequency, and dwell time are proxies for the allocation of attention in neurocognitive and aviation tasks[4,5,6]. Highly successful performers and experts show significant fixation biases toward task-critical areas compared to less successful performers or novices[4,7,8]. Spatiotemporal aspects of gaze are captured through changes in dwell time patterns across various areas of interest (AOIs) or measures of fixation distribution (i.e., Stationary Gaze Entropy: SGE). In the context of laboratory-based paradigms, average fixation duration, scan path length, and gaze sequence complexity (i.e., Gaze Transition Entropy: GTE) tend to increase due to the increased scanning and processing required to problem-solve and elaborate on more

challenging task goals/solutions[4,7].

Conversely, aviation studies have demonstrated that scan path length and gaze sequence complexity decrease with task complexity and cognitive load. This discrepancy highlights the fact that understanding the task components and the demands of the paradigm being employed is critical for the accurate interpretation of gaze metrics[2]. Altogether, research to date supports that gaze measures provide meaningful, objective insight into information processing that underlies the differences in task difficulty, cognitive load, and task performance. With advances in eye tracking technology, examining gaze behavior in 'the wild' is an emerging area of research with tangible applications toward advancing occupational training in the fields of medicine[9,10,11] and aviation[12,13,14].

The current work aims to further examine the utility of using gaze-based metrics to gain insight into information processing by specifically employing a wearable eye tracker during a flight simulation task in low-time pilots. This study expands on previous work that used a head-stabilized eye tracker (i.e., EyeLink II) to examine differences in gaze behavior metrics as a function of flight difficulty (i.e., changes in weather conditions)[5]. Specifically, the current work examined traditional and dynamic gaze behavior metrics while task difficulty was manipulated via the introduction of an in-flight emergency scenario (i.e., unexpected total engine failure). The gaze metric analysis and the introduction of an in-flight emergency scenario will help to determine how changes in gaze behavior underlie changes in information processing during more challenging task conditions. In addition, they will showcase the utility of the wearable eye tracker in capturing natural head and eye movements along with the use of a computer vision-based algorithm to conduct AOI-based analyses in real-world environments.

In line with previous work[5,8], it was hypothesized that an increase in task difficulty would result in a reduction in aircraft control and landing success, a focusing of gaze toward fewer, task-critical AOIs, which would be evident through the increase in dwell time toward select AOIs (i.e., the front window), and a reduction in SGE and GTE.

PROTOCOL:

The following protocol can be applied to various experiments involving a wearable eye-tracker and a flight simulator but details an experiment in which eye-tracking data were collected alongside complex aviation-related tasks in a flight simulator (see **Table of Materials**). The simulator was configured to be representative of a Cessna 172 and was used with the necessary instrument panel (steam gauge configuration), an avionics/GPS system, an audio/lights panel, a breaker panel, and a Flight Control Unit (FCU) (see **Figure 1**). This study was approved by the University of Waterloo's Office of Research Ethics (43564; Date: Nov 17, 2021). All participants (N = 24; 14 males, 10 females; mean age = 22 years; flight hours range: 51–280 h) provided written informed consent.

(Place figure 1 here)

1. Participant screening and informed consent

1.1. Screen the participant via a self-report questionnaire based on inclusion/exclusion criteria[2,5]: possession of at least a Private Pilot License (PPL), normal or corrected-to-normal vision, and no previous diagnosis with a neuropsychiatric/neurological disorder or learning disability.

1.2. Inform the participant about the study objectives and procedures through a detailed briefing handled by the experimenter and the supervising flight instructor/simulator technician. Review the risks outlined in the institution's ethics review board-approved consent document. Answer any questions about the potential risks. Obtain written informed consent before beginning any study procedures.

2. Hardware/software requirements and start-up

2.1. Flight simulator (typically completed by the simulator technician)

2.1.1. Turn the simulator and projector screens on. If one of the projectors does not turn on at the same time as the others, restart the simulator.

2.1.2. On the **Instruction Screen**, press the **Presets** tab on the screen and verify that the **Position** and/or **Weather** presets required are available. If needed, create a new type of preset; consult the technician for help.

2.2. Collection laptop

2.2.1. Turn the laptop and wearable eye-tracker on and log in with credentials.

2.2.2. When prompted, either select a preexisting profile or make one if testing a new participant. Alternatively, select the **Guest** option to overwrite its last calibration.

2.2.2.1. To create a new profile, scroll to the end of the profile list and click **Add**.

2.2.3. Set the **profile ID** to the **participant ID**. This profile ID will be used to tag the folder, which holds eye tracking data after a recording is complete.

2.2.4. Select the nosepiece.

2.3. Glasses calibration

NOTE: The glasses must stay connected to the laptop to record. Calibration with the box only needs to be completed once at the start of data collection.

2.3.1. Open the eye tracker case and take out the glasses.

2.3.2. Connect the USB to micro-USB cable from the laptop to the glasses. If prompted on the laptop, update the firmware.

2.3.3. Locate the black calibration box inside the eye tracker case.

2.3.4. On the collection laptop, in the **eye tracking Hub**, choose **Tools | Device Calibration**.

2.3.5. Place the glasses inside the box and press **Start** on the popup window to begin calibration.

2.3.6. Remove the glasses from the box once calibration is complete.

2.4. Nosepiece fit

2.4.1. Instruct the participant to sit in the cockpit and put on the glasses.

2.4.2. In **eye tracking Hub**, navigate to **File | Settings | Nose Wizard**.

2.4.3. Check the **adjust fit of your glasses** box on the left side of the screen. If the fit is **Excellent**, proceed to the next step. Otherwise, click the box.

2.4.4. Tell the participant to follow the fit recommendation instructions shown onscreen: set the nosepiece, adjust the glasses to sit comfortably, and look straight ahead at the laptop.

2.4.5. If required, swap out the nosepiece. Pinch the nosepiece in its middle, slide it out from the glasses, and then slide another one in. Continue to test the different nosepieces until the one that fits the participant the best is identified.

2.5. Air Traffic Control (ATC) calls

NOTE: If the study requires ATC calls, have the participant bring in their own headset or use the lab headset. Complete eyeball calibration only after the participant puts on the headset as the headset can move the glasses on the head, which affects calibration accuracy.

2.5.1. Check that the headset is hooked up to the jack on the left underside of the instrument panel.

2.5.2. Instruct the participant to put the headset on. Ask them not to touch it or take it off until the recording is finished.

NOTE: Re-calibration is required each time the headset (and therefore the glasses) is moved.

2.5.3. Do a radio check.

2.6. Eyeball calibration

NOTE: Whenever the participant shifts the glasses on their head, they must repeat eyeball calibration. Ask the participant not to touch the glasses until their trials are over.

2.6.1. In **eye tracking Hub**, navigate to the **parameter box** at the left of the screen.

2.6.1.1. Check the **calibration mode** and choose **fixed gaze** or **fixed head** accordingly.

2.6.1.2. Check that the calibration points are a **5 x 5 grid**, for **25 points total**.

2.6.1.3. Check the **validation mode** and make sure it matches the **calibration mode**.

2.6.1.4. Examine the eye tracking outputs and verify that everything that needs to be recorded for the study is checked using the tick boxes.

2.6.2. Click **File | Settings | Advanced** and check that the **sampling rate** is **250 Hz**.

2.6.3. Check the **Calibrate your eye tracking** box on the screen using the mouse. The calibration instructions will vary based on the mode. To follow the current study, use the **fixed gaze** calibration mode: instruct the participants to move their head so that the box overlaps with the black square and they align. Then, ask the participant to focus their gaze on the crosshair in the black square and press the space bar.

2.6.4. Press the **Validate your setup** box. The instructions will be the same as step 2.6.3. Check that the **validation figure MAE** (Mean Absolute Error) is **<1°**. If not, then repeat steps 2.6.3 and 2.6.4.

2.6.5. Press **Save Calibration** to save the calibration to the profile each time calibration and validation are completed.

2.7. iPad usage

NOTE: The iPad is located to the left of the instrument panel (see **Figure 1**). It is used for questionnaires typically after flight.

2.7.1. Turn the iPad on and make sure it is connected to the Internet.

2.7.2. Open a window in Safari and enter the link for the study questionnaire.

3. Data collection

NOTE: Repeat these steps for each trial. It is recommended that the laptop is placed on the bench outside the cockpit.

3.1. On the flight simulator computer, in the **instruction** screen, press **Presets**, and then choose the desired **Position** preset to be simulated. Press the **Apply** button and watch the screens surrounding the simulator to verify that the change happens.

3.2. Repeat step 3.1 to apply the **Weather** preset.

3.3. Give the participant any specific instructions about the trial or their flight path. This includes telling them to change any settings on the instrument panel before they begin.

3.4. In the instruction screen, press the orange **STOPPED** button to start data collection. The color will change to green and the text will say **FLYING**. Be sure to give a verbal cue to the participant so that they know they can start flying the aircraft. The recommended cue is "**3, 2, 1, you have controls**" as the orange **stop** button is pressed.

3.5. In the collection laptop, press **Start Recording** so that the eye tracker data are synced with the flight simulator data.

3.6. When the participant has completed their circuit and landed, wait for the aircraft to stop moving.

NOTE: It is important to wait because during postprocessing; data are truncated when the groundspeed settles at 0. This gives consistency for the endpoint of all trials.

3.7. In the instruction screen, press the green **FLYING** button. The color will return to orange and the text will say **STOPPED**. Give a verbal cue during this step when data collection is about to end. The recommended cue is "**3, 2, 1, stop**".

3.8. Instruct the participant to complete the posttrial questionnaire(s) on the iPad. Refresh the page for the next trial.

NOTE: The current study used the Situation Awareness Rating Technique (SART) self-rating questionnaire as the only posttrial questionnaire[17].

4. Data processing and analysis

4.1. Flight simulator data

NOTE: The .csv file copied from the flight simulator contains more than 1,000 parameters that can be controlled in the simulator. The main performance measures of interest are listed and described in **Table 1**.
(Place Table 1 here)

4.1.1. For each participant, calculate the **success rate** using Eq (**1**). by taking the percentage across the task conditions. Failed trials are identified by predetermined criteria programmed within the simulator that terminates the trial automatically on touchdown due to plane orientation and vertical speed. Carry out posttrial verification to ensure that this criterion was aligned with actual aircraft limitations (i.e., Cessna 172 landing gear damage/crash is evident at vertical speeds > 700 feet/min [fpm] upon touchdown).

$$\text{Success rate} = \frac{Total\ successful\ landing\ trials}{Total\ landing\ trials} \times 100 \qquad (\mathbf{1})$$

NOTE: Lower success rate values indicate worse outcomes as they are associated with a reduction in successful landing attempts.

4.1.2. For each trial, calculate **completion time** based on the timestamp associated with the plane coming to a complete stop on the runway (i.e., GroundSpeed = 0 knots).

NOTE: A shorter completion time may not always equate to better performance. Caution must be taken to understand how the task conditions (i.e., additional winds, emergency scenarios, etc.) are expected to impact completion time.

4.1.3. For each trial, determine **landing hardness** based on aircraft vertical speed (fpm) at the time the aircraft initially touches down on the runway. Ensure that this value is taken at the same timestamp associated with the first change in **AircraftOnGround** status from **0** (**in air**) to **1** (**on ground**).

NOTE: Values within the range of **-700 fpm** to **0 fpm** are considered **safe**, with values closer to 0 representing softer landings (i.e., better). **Negative values** represent **downward vertical speed**; **positive values** represent **upward vertical speed**.

4.1.4. For each trial, calculate the **landing error (°)** based on the difference between the touchdown coordinates and the reference point on the runway (center of the 500 ft markers). Using the reference point, calculate the landing error using Eq (**2**).

$$\text{Difference} = \sqrt{((\Delta\ \text{Latitude})^2 + (\Delta\ \text{Longitude})^2)} \qquad (\mathbf{2})$$

NOTE: Values below 1° are shown to be $\text{normal}_{5,15}$. Large values indicate a larger landing error associated with aircraft touchdown points that are farther away from the landing zone.

(Place figure 2 here)

4.1.5. Calculate the **means** across all participants for each **performance outcome variable** for each **task condition**. Report these values.

4.2. Situation awareness data

4.2.1. For each trial, calculate the SA score based on the self-reported SART scores across the 10 dimensions of SA[17].

4.2.1.1. Use the SART questionnaire[17] to gauge the participants' subjective opinion on various domains related to task difficulty, as well as the supply and demand attentional resources required during task performance.

4.2.1.2. Using a 7 point Likert scale, ask the participants to rate their perceived experience on probing questions, including the complexity of the situation, division of attention, spare mental capacity, and information quantity and quality.

4.2.1.3. Combine these scales into larger dimensions of attentional demands (Demand), attentional supply (Supply), and situation understanding (Understanding).

4.2.1.4. Use these ratings to calculate a measure of SA based on Eq (**3**):

SA = Understanding – (Demand-Supply) (**3**)

NOTE: Higher scores on scales combined to provide a measure of understanding suggest the participant has a good understanding of the task at hand. Similarly, high scores in the supply domain suggest that the participant has a significant amount attentional resources to devote toward a given task. In contrast, a high demand score suggests that the task requires a significant amount of attentional resources to complete. It is important to clarify that these scores are best interpreted when compared across conditions (i.e., easy versus difficult conditions) instead of used as stand-alone measures.

4.2.2. Once data collection is completed, calculate the means across all participants for each task condition (i.e., basic, emergency). Report these values.

4.3. Eye tracking data

4.3.1. Use an eye tracking batch script for manually defining AOIs for use in gaze mapping. The script will open a new window for key frame selection that should clearly display all key AOIs that will be analyzed. Scroll through the video and choose a frame that shows all the AOIs clearly.

(Place Figure 3 here)

4.3.2. Following the instructions onscreen, draw a rectangle over a region of the frame that will be visible throughout the whole video, unique, and remain stable.

(Place Figure 4 here)

NOTE: The purpose of this step is to generate an “in-screen” coordinate frame that can be used across the video recording since head movements result in the location of objects in the environment changing over the course of the video recording.

4.3.3. Draw a rectangle for each AOI in the picture, one at a time. Name them accordingly. Click **Add more** to add a new AOI and press **Done** on the last one. If the gaze coordinates during a given fixation land within the object space as defined in the “in-screen” coordinate frame, label that fixation with the respective AOI label.

(Place Figure 5 here)

NOTE: The purpose of this step is to generate a library of object coordinates that are then used as references when comparing gaze coordinates to in-screen coordinates.
There are typically 10 AOIs, but this depends on how the flight simulator is configured. The instrument panel may be different. In line with previous work[5,15], the current study uses the following **AOIs**: **Airspeed, Attitude, Altimeter, Turn Coordinator, Heading, Vertical Speed, Power, Front Window, Left Window**, and **Right Window** (see **Figure 1**).

4.3.4. Let the script start processing the AOIs and generate fixation data. It generates a plot showing the saccades and fixations over the video.

(Place Figure 6 here)

4.3.5. Two new files will be created: fixations.csv and aoi_parameters.yaml. The batch processor will complete the postprocessing of gaze data for each trial and each participant.

NOTE: The main eye tracking measures of interest are listed in **Table 2** and are calculated for each AOI for each trial.

(Place Table 2 here)

4.3.6. For each trial, calculate the **traditional gaze metrics[4,5]** for each AOI based on the data generated in the fixation.csv file.

NOTE: Here, we focus on **dwell time (%)**, which is calculated by dividing the sum of fixations for a particular AOI by the sum of all fixations and multiplying the quotient by 100 to get the percentage of time spent in a specific AOI. There is no inherent negative/positive interpretation from the calculated dwell times. They give an indication of where attention is predominantly being allocated. Longer average fixation durations are indicative of increased processing demands.

4.3.7. For each trial, calculate the **blink rate** using Eq (**4**):

Blink rate = Total blinks/completion time (**4**)

NOTE: Previous work has shown that blink rate is inversely related to cognitive load[2,6,13,18,19].

4.3.8. For each trial, calculate the SGE using eq (**5**)[20]:

$$H_{SGE}(V) = -\sum_{v \in V} v \cdot log(v) \quad (\mathbf{5})$$

Where *v* is the probability of viewing the i[th] AOI and *V* is the number of AOIs.

NOTE: Higher SGE values are associated with a larger fixation dispersion, whereas lower values are indicative of a more focalized allocation of fixations[21].

4.3.9. For each trial, calculate the GTE using eq (**6**)[22]:

$$H_{GTE}(M) = -\sum_{i=1}^{10} V_i \sum_{j=1}^{10} M_{i,j} \cdot log\left(M_{i,j}\right) \quad (\mathbf{6})$$

Where *V* is the probability of viewing the i[th] AOI, and *M* is the probability of viewing the j[th] AOI given the previous viewing of the i[th] AOI.

NOTE: Higher GTE values are associated with more unpredictable, complex visual scan paths, whereas lower GTE values are indicative of more predictable, routine visual scan paths.

4.3.10. Calculate the **means** across all participants for each eye tracking output variable (and AOI when indicated) and each task condition. Report these values.

REPRESENTATIVE RESULTS:

The impact of task demands on flight performance

The data were analyzed based on successful landing trials across basic and emergency conditions. All measures were subjected to a paired-samples' *t*-test (within-subject factor: task condition [basic, emergency]). All *t*-tests were performed with an alpha level set at 0.05. Four participants crashed during the emergency scenario trial and were not included in the main analyses. All variables except for success rate exclusively examine successful trials.

Success rate (%) produced a main effect of condition, $t(23) = 2.145$, $p = 0.043$. Specifically, emergency trials (mean = 83%) resulted in significantly more failed landings (i.e., crashed) than basic trials (mean = 100%) (**Figure 7A**). Completion time (s) produced a main effect of condition, $t(19) = 8.420$, $p < 0.001$. Emergency trials were completed significantly more quickly (mean = 121 s, SD = 3.9) than basic trials (mean = 174 s, SD = 5.6) (**Figure 7B**). Landing error produced a main effect of condition, $t(19) = -2.669$, $p = 0.015$, $\eta_p^2 = 0.242$. Specifically, basic trials were associated with significantly lower landing error (i.e., higher landing accuracy) (mean = 0.046 °, SD = 0.010) compared to emergency trials (mean = 0.216°, SD = 0.070) (**Figure 7C**), which included one person who landed in the field next to the runway. Lastly, landing hardness, did not significantly change between conditions ($p = 0.062$) (**Figure 7D**).

(Place Figure 7 here)

As task difficulty increased, performance was negatively impacted. This was mostly evident through the reduction in landing success rate, increased landing error, as well as reduced subjective SA. Note that the reduction in landing completion time was more a consequence of insufficient power to the engine from the engine failure, resulting in a necessary change in flight trajectory that reduced the flight path significantly so that the aircraft could land on the runway safely.

The impact of task demands on situation awareness

Subjective scores from the SART questionnaire produced a general SA score, which demonstrated a main effect of condition, $t(19) = 9.148$, $p < 0.001$. Specifically, subjective SA scores were lower for emergency trials (mean = 11.7, SD = 1.4) compared to basic trials (mean = 21.4, SD = 1.2) (**Figure 8A**). A closer examination of the SART questionnaire subcomponents revealed that SA supply, SA demand, and SA understanding all yielded a main effect of condition (**Figure 8B-D**). Specifically, SA supply increased significantly from the basic condition (mean = 18.7, SD = 0.8) to the emergency condition (mean = 21.9, SD = 0.9), $t(19) = -4.921$, $p < 0.001$. Similarly, SA demand increased significantly from the basic condition (mean = 8.1, SD = 1.5) to the emergency condition (mean = 19.3, SD = 1.8), $t(19) = -10.696$, $p < 0.001$. Lastly, SA understanding decreased significantly from the basic condition (mean = 10.7, SD = 0.4) to the emergency condition (mean = 9.0, SD = 0.6), $t(19) = 3.187$, $p = 0.005$.

(Place Figure 8 here)

The impact of task demands on gaze behavior

Traditional gaze metrics

Dwell time (%) across several AOIs demonstrated a main effect of condition (**Figure 9**). The majority of AOIs demonstrated a decrease in dwell time from the basic condition to the emergency condition including, attitude, $t(19) = 2.322$, $p = 0.031$, altimeter, $t(19) = 2.822$, $p = 0.011$, turn coordinator, $t(19) = 2.698$, $p = 0.014$, heading, $t(19) = 2.175$, $p = 0.042$, vertical speed, $t(19) = 2.357$, $p = 0.029$, and power gauges, $t(19) = 3.036$, $p = 0.007$. In contrast, the airspeed indicator time increased from the basic condition to the emergency condition, $t(19) = -2.376$, $p = 0.029$. All other AOIs were not significantly modulated by condition (all p values > 0.165). Dwell time means and standard deviations for all AOIs are shown in **Table 3**.

(Place Figure 9 here)
(Place Table 3 here)

Blink rate demonstrated a main effect of condition, $t(19) = -2.713$, $p = 0.014$ (**Figure 10**). Specifically, the blink rate increased significantly from the basic condition (mean = 0.354 blinks/sec, SD = 0.192) to the emergency condition (mean= 0.460 blinks/sec; SD= 0.285).

(Place Figure 10 here)

Advanced gaze metrics

SGE, GTE, and normalized means and standard deviations for all AOIs across conditions are reported in **Table 4**. SGE and GTE demonstrated a significant reduction between basic and emergency task conditions (**Figure 11**), $t(19) = 4.833$ and 4.833, $ps < 0.001$, respectively.

(Place Figure 11 here)
(Place Table 4 here)

Taken together, significant changes in gaze behavior were also reported in parallel with the noted decrements to performance and SA. Specifically, responding to the emergency was associated with a significant reduction in the allocation of attention towards several AOIs, which was evident through the reduction in SGE (i.e., fixation dispersion) and shorter dwell time in a number of AOIs (i.e., attitude, altimeter, turn coordinator, heading, vertical speed, and power indicators). On the other hand, the airspeed indicator received significantly more attention as this became a prime source of information for managing the emergency landing condition; a necessary indicator that helps to establish optimal glide speed to prevent the plane from stalling[27]. These changes in attention allocation were also associated with reduced scanning complexity (GTE), which indicates that participants adopted a strategy by shifting their attention toward fewer AOIs in a more routine/predictable manner. During emergency scenarios, these changes in gaze behavior are to be expected since the protocol for managing such an event requires directing attention mainly to the airspeed gauge and the runway (i.e., front/left window). These data confirm that the task manipulation was difficult enough to impact task performance and the supporting information processing mechanisms. More importantly, the gaze behavior findings provide additional empirical evidence that the scanning of information across cockpit AOIs is reduced during challenging task conditions at the cost of spending more time on AOIs that have higher relevance for decision-making and problem-solving[5,8,28,29]. This is a critical finding that may suggest that the emergency scenario is a well-learned task as previous work using highly uncertain events during flight resulted in higher exploratory activity (i.e., higher SGE/GTE) when the tasks were not well practiced or trained[30]. Nevertheless, more research is required to determine how gaze behavior, and, thus, information processing change in response to these task manipulations in novice pilots or during failed landing attempts where the observed changes in gaze behavior may not be present. Such a finding may indicate that a lack of selective attention is associated with poorer task performance during uncertain/unexpected events.

FIGURE AND TABLE LEGENDS:

Figure 1: Flight simulator environment. An illustration of the flight simulator environment. The participant's point of view of the cockpit replicated that of a pilot flying a Cessna 172, preset for a downwind-to-base-to-final approach to Waterloo International Airport, Breslau, Ontario, CA. The orange boxes represent the ten main areas of interest used in the gaze analyses. These include the (1) airspeed, (2) attitude, (3) altimeter, (4) turn coordinator, (5) heading, (6) vertical speed, and (7) power indicators, as well as the (8) front, (9) left, and (10) right windows.

Figure 2: Landing scenario flight path. Schematic of (**A**) the landing circuit completed in all trials

and (**B**) the runway with the 500 ft markers that were used as the reference point for the landing zone (i.e., center orange circle).

Figure 3: Area of Interest mapping. An illustration of the batch script demonstrating a window for frame selection. The selection of an optimal frame involves choosing a video frame that includes most or all areas of interest to be mapped.

Figure 4: Generating Area of Interest mapping "in-screen" coordinates. An illustration of the batch script demonstrating a window for "in-screen" coordinates selection. This step involves the selection of a square/rectangular region that remains visible throughout the recording, is unique to the image, and remains static.

Figure 5: Identifying Area of Interest to be mapped. An illustration of the batch script window that allows for the selection and labelling of areas of interest. Abbreviation: AOIs = areas of interest.

Figure 6: Batch script processing. An illustration of the batch script processing the video and gaze mapping the fixations made throughout the trial.

Figure 7: Aircraft performance. Individual data points and their respective group means for (**A**) success rate (%), (**B**) completion time (s), (**C**) landing error (°), and (**D**) landing hardness (fpm) are demonstrated for basic and emergency conditions. Success rate and completion time decreased in the emergency condition relative to the basic condition. Landing error increased in the emergency condition relative to the basic condition. Landing hardness was not significantly different between conditions. Statistical test used: paired-samples *t*-test. Error bars represent SEM. *$p \leq 0.05$, **$p \leq 0.01$, ***$p \leq 0.001$. Abbreviation: fpm = feet/min.

Figure 8: Situation awareness scores. Individual data points and their respective group means for (**A**) situational awareness, (**B**) SA Supply, (**C**) SA Demand, and (**D**) SA Understanding are demonstrated for basic and emergency conditions. Situation awareness and SA understanding decreased in the emergency condition relative to the basic condition. SA supply and demand increased in the emergency condition relative to the basic condition. Statistical test used: paired-samples *t*-test. Error bars represent SEM. *$p \leq 0.05$, **$p \leq 0.01$, ***$p \leq 0.001$.

Figure 9: Dwell time. Group means for dwell time percentage changes (emergency - basic) are demonstrated for all 10 areas of interest. Attitude, altimeter, turn coordinator, heading, vertical speed, and power gauges decreased in the emergency condition relative to the basic condition. The airspeed indicator increased in the emergency condition relative to the basic condition. Statistical test used: paired-samples *t*-test. Error bars represent SEM. *$p \leq 0.05$, **$p \leq 0.01$, ***$p \leq 0.001$.

Figure 10: Blink rate. Individual data points and their respective group means for blink rate (blinks/sec) are demonstrated for basic and emergency conditions. Blink rate increases in the emergency condition relative to the basic condition. Statistical test used: paired-samples *t*-test.

Error bars represent SEM. *$p \leq 0.05$, **$p \leq 0.01$, ***$p \leq 0.001$.

Figure 11: Entropy metrics. Individual data points and their respective group means for (**A**) stationary gaze entropy and (**B**) gaze transition entropy are demonstrated for basic and emergency conditions. Both stationary gaze entropy and gaze transition entropy decrease in the emergency condition relative to the basic condition. Statistical test used: paired-samples *t*-test. Error bars represent SEM. *$p \leq 0.05$, **$p \leq 0.01$, ***$p \leq 0.001$. Abbreviations: SGE = stationary gaze entropy; GTE = gaze transition entropy.

Table 1: Simulator performance outcome variables. Aircraft performance-dependent variables and their definitions.

Table 2: Eye tracking outcome variables. Gaze behavior-dependent variables and their definitions.

Table 3: Dwell time (%) values by condition. Mean (standard deviation) dwell time (%) values across all areas of interest and task conditions (i.e., basic, emergency).

Table 4: Entropy values by condition. Mean (standard deviation) for all entropy (bits) values across all task conditions (i.e., basic, emergency).

DISCUSSION:

The eye tracking method described here enables the assessment of information processing in a flight simulator environment via a wearable eye tracker. Assessing the spatial and temporal characteristics of gaze behavior provides insight into human information processing, which has been studied extensively using highly controlled laboratory paradigms[4,7,23]. Harnessing recent advances in technology allows the generalization of eye tracking research to realistic paradigms with higher fidelity, thereby mimicking more naturalistic settings. The aim of the current study was to characterize the effects of task difficulty on gaze behavior, flight performance, and SA during a simulated landing scenario in low-time pilots. Participants were asked to perform the landing task in high-visibility, visual flight rules (VFR) conditions that included a basic landing scenario and a challenging manipulation, which was an unexpected introduction of an in-flight emergency (i.e., total engine failure). In line with our hypothesis, the introduction of an emergency during the landing scenario resulted in notable performance and SA decrements along with significant changes to gaze behavior. These methods, results, and the insight gained through gaze behavior analysis about information processing, flight performance, and SA are further discussed below.

With respect to the methodology, there are several critical points to be kept in mind including, modifications of the protocol and significance of the protocol with respect to the literature. The eye-tracking data were collected alongside complex aviation-related tasks in an ALSIM AL-250 flight simulator. The ALSIM simulator was configured to be representative of a Cessna 172 and was used with the necessary instrument panel (steam gauge configuration), an avionics/GPS system, an audio/lights panel, a breaker panel, and a Flight Control Unit (FCU) (see **Figure 1**). The

simulated aircraft functioned at a sampling frequency of 30 Hz and was controlled with a yoke, throttle lever, and rudder pedals. In this protocol, a wearable eye tracker was used to assess changes in gaze behavior (i.e., a proxy for information processing) during the performance of the basic landing scenario (high visibility [>20 miles], low winds [0 knots]), and an emergency landing scenario (high visibility [>20 miles], low winds [0 knots], unexpected total engine failure).

The current experiment required participants to complete a total of five landing trials: four basic trials and one emergency trial. The sequence of these trials remained the same across participants to ensure that the introduction of the in-flight emergency did not impact the natural gaze behaviors being captured during the other trials, which included other manipulations that were part of a larger experiment$_2$. In addition to the pilot briefing at the beginning of the session, participants were provided two practice trials before the experimental trials for cockpit familiarization. The AdHawk MindLink eye tracker (250 Hz, <2° spatial resolution, front-facing camera) was used in the current study but can be replaced by any high-quality, commercially available, wearable eye tracker. However, the wearable eye tracker should have a spatial resolution of <2°, a front facing camera, and a sampling rate of 120 Hz to properly identify eye movement characteristics related to saccade/fixation event detection[16].

Based on previous work by Ayala and colleagues$_2$, the front-facing camera is required to provide video recordings of the task environment, which is then used for gaze mapping by superimposing the eye-tracking gaze coordinates. These coordinates are then compared to the manually defined coordinates for task-relevant AOIs within the gaze mapping algorithm. The ALSIM flight simulator can also be replaced with other forms of flight simulation (i.e., PC-based flight simulator, VR flight simulator)[5]. It is, however, important to note that the mapping of AOIs should be similar across these environments to safeguard against the risk of discordant AOI mapping becoming a confounding variable in the gaze results. Moreover, the mapping used in the current work was based on the knowledge that these predefined spaces are contextually relevant to the pilot, who must focus their gaze on these critical sources of information for safe and successful aircraft operation. The selected eye tracker and simulation environments were specifically chosen for this work because they provided an accurate assessment of gaze behavior in a high-fidelity, ecological environment.

As for the results, the current work showcases the characterization of gaze behavior using a wearable eye tracker that leverages the use of a computer vision algorithm to aid in gaze mapping in a high-fidelity, immersive flight simulation environment. The application of this method provides several advantages to the assessment of gaze behavior and general information processing changes in naturalistic settings, which are outlined here. First, the protocol moves beyond the standard evaluation of eye movements in a lab-based 2D screen environment by capturing combined head and eye movements in an immersive 3D space[2,4,5]. Second, most studies done in 3D space have lacked the capacity to conduct efficient gaze analyses based on defined AOIs[10,24,25]. This is highly pertinent to domain-specific tasks in everyday life as the AOI-based analyses conducted here provide critical context that is necessary for the proper interpretation of the results. The current study collected behavioral performance data, subjective SA, and eye tracking data to provide a comprehensive assessment of how each data stream

relates back to underlying cognitive functions such as attention and decision-making. Using this type of multimodal analysis framework, several studies have demonstrated that the amount of time for which the gaze is directed toward particular AOIs is biased toward task-critical regions that require more information-processing resources, which, in turn, is associated with successful task performance[4,5,7,26]. Similar to the findings noted in the current work, decrements in performance associated with increases in task difficulty in neurocognitive tasks (i.e., increased completion time, increased planning time, decreased task accuracy) were all shown to be associated with a significant increase in the focusing of attentional resources (i.e., specific increases in dwell time and fixation durations, and decreased SGE/GTE) toward informationally dense AOIs that were important for successful task completion[4,7].

Users should also consider several critical elements of the described eye tracking protocol. First, it is well known that eye tracking is a method that indirectly examines changes in information processing through overt shifts in attention that are captured by gaze behavior metrics. The current method is, therefore, limited in the extent to which it can identify and examine covert processes that may not be associated with explicit shifts in gaze but may be relevant for performance proficiency[31]. Second, the current work shows that there is still much work to be done to understand how some metrics such as blink rate truly relate back to aspects of human cognition and action. Specifically, previous work has suggested that blink rate is inversely correlated with task difficulty[2,6,13,18,19]. However, the current work provides contradictory evidence for this as blink rate was shown to increase in the emergency scenario (i.e., an increase in task difficulty). Given the lack of consensus regarding the interpretation of blinks as a proxy measure of task difficulty or cognitive load, this is an area that requires further investigation to understand the mechanisms underlying blinks and why they change at all when task demands are altered. Exploring blink rate is important because this signal is relatively easy to measure with eye tracking; however, the utility and application to more complex real-world tasks need to be further established. Third, the current paradigm used a fixed condition schedule that resulted in the emergency trial always being completed last. This was mainly a consequence of the data collection being a part of a larger study that examined changes in information processing across various environmental conditions. The emergency landing scenario could have altered the normal gaze behavior during these alternative flight conditions. Thus, to prevent any within-session alterations to gaze behavior due to the early introduction of an uncertain event, the emergency scenario was completed at the end of the experimental session. Although it could be argued that a lack of randomization in the trial sequence could result in a practice effect, the results outlined in the current work show that this is clearly not the case as performance and SA decreased, while significant changes to gaze became evident only in the last emergency trial and not in the preceding four basic trials. Lastly, the lack of synchronicity across collection devices (i.e., eye tracker and flight simulator) is a lingering issue with respect to the alignment of data streams for flight performance and eye tracking data. Although the current work attempted to have synchronization between the two devices by pressing the start/recording buttons at the same time, there was always the chance that human error and normal motor control variability in button presses would become sources of synchronization errors that should be considered. This limited the extent to which temporal contingencies between specific eye movements and actions could be examined with the current method and is a technological shortfall that requires

further development.

Typical applications of eye tracking methods have linked changes in information processing to several health disorders (i.e., traumatic brain injury, Schizophrenia, Parkinson's Disease)[7,32,33] and have used gaze behavior to shed insight on the development of assessment and training methods. The latter applications aim to use eye tracking to enhance the type of feedback received by trainees across several domains (i.e., medicine and aviation)[2,10]. The improved capability of wearable eye trackers and AOI-based analyses enhances the contextual information provided to researchers and industry applications alike. For instance, in high-fidelity environments, the enhanced analysis allows for the accurate identification of what individuals are overtly attending to, when, and how effectively information is being processed[2,15]. However, the utility of using the method as a means to identify skill proficiency or level of expertise in natural environments is yet to be extensively explored.

One significant limitation to training and assessment applications are the need for larger scale studies across various levels of flight experience to establish the reliability and normative ranges, as has been done with simpler eye movements in highly controlled laboratory experiments for basic neurological and cognitive function[34,35]. Furthermore, industry training and assessment applications would benefit from an enhanced understanding of how gaze is distributed through the task environment and how it changes during task performance, particularly in superior performers. One exciting prospect is combining eye tracking with motion capture systems, which can be used to quantify the motor actions generated during a given task. For instance, movement variability has been shown to be an indicator of skill development[36]. The alignment of motor control data with eye tracking data may provide greater support for the assessment of skill proficiency throughout training. As such, this provides a unique opportunity to develop a comprehensive analytical method to assess and gain insight into the dynamic interactions between perceptual, cognitive, and motor processes underlying human performance and learning.

This study investigated the utility of a wearable eye tracker and gaze mapping algorithm in characterizing gaze behavior during a simulated flight task. The eye tracking method was useful in quantifying changes in gaze behavior, which provided insight into changes in information processing associated with responding to and managing an in-flight emergency. Gaze measures indicated a distinct change in attention allocation. Specifically, dwell time, SGE, and GTE demonstrated a focusing of attention toward fewer AOIs with higher problem relevance. Notably, the gaze mapping algorithm and wearable eye tracker are relatively nascent technology and, thus, should be used and further developed in future work, given the known limitations in synchronization across multiple hardware collection devices.

ACKNOWLEDGMENTS:
This work is supported in part by the Canadian Graduate Scholarship (CGS) from the Natural Sciences and Engineering Research Council (NSERC) of Canada, and the Exploration Grant (00753) from the New Frontiers in Research Fund. Any opinions, findings, and conclusions or recommendations expressed in this material are the author(s) and do not necessarily reflect

those of the sponsors.

DISCLOSURES:
No competing financial interests exist.

REFERENCES:

1. de Brouwer, A. J., Flanagan, J. R., Spering, M. Functional use of eye movements for an acting system. *Trends Cogn Sci.* **25** (3), 252-263 (2021).
2. Ayala, N., Kearns, S., Cao, S., Irving, E., & Niechwiej-Szwedo, E. Investigating the role of flight phase and task difficulty on low-time pilot performance, gaze dynamics and subjective situation awareness during simulated flight. *Journal of Eye Movement Research*. ***17*** (1), (2024a).
3. Land, M. F., & Hayhoe, M. In what ways do eye movements contribute to everyday activities?. *Vision Research*, ***41***(25-26), 3559-3565 (2001).
4. Ayala, N., Zafar, A., & Niechwiej-Szwedo, E. Gaze behavior: a window into distinct cognitive processes revealed by the Tower of London test. *Vision Research*. ***199****, 108072* (2022).
5. Ayala, N., Zafar, A., Kearns, S., Irving, E., Cao, S., & Niechwiej-Szwedo, E. The effects of task difficulty on gaze behavior during landing with visual flight rules in low-time pilots. *Journal of Eye Movement Research*, ***16***(1), (2023).
6. Glaholt, M. G. Eye tracking in the cockpit: a review of the relationships between eye movements and the aviators cognitive state. https://apps.dtic.mil/sti/pdfs/AD1000097.pdf (2014).
7. Hodgson, T. L., Bajwa, A., Owen, A. M., & Kennard, C. The strategic control of gaze direction in the Tower-of-London task. *Journal of Cognitive Neuroscience*, ***12***(5), 894–907 (2000).
8. van De Merwe, K., Van Dijk, H., & Zon, R. Eye movements as an indicator of situation awareness in a flight simulator experiment. *The International Journal of Aviation Psychology*, ***22***(1), 78-95 (2012).
9. Kok, E. M., & Jarodzka, H. Before your very eyes: The value and limitations of eye tracking in medical education. *Medical Education*, ***51***(1), 114-122 (2017).
10. Di Stasi, L. L., Diaz-Piedra, C., Rieiro, H., Sanchez Carrion, J. M., Martin Berrido, M., Olivares, G., & Catena, A. Gaze entropy reflects surgical task load. *Surgical Endoscopy*, ***30***, 5034-5043 (2016).
11. Laubrock, J., Krutz, A., Nübel, J., & Spethmann, S. Gaze patterns reflect and predict expertise in dynamic echocardiographic imaging. *Journal of Medical Imaging*, ***10***(S1), S11906-S11906 (2023).
12. Brams, S., Hooge, I. T., Ziv, G., Dauwe, S., Evens, K., De Wolf, T., ... & Helsen, W. F. Does effective gaze behavior lead to enhanced performance in a complex error-detection cockpit task?. *PloS one*, ***13***(11), e0207439 (2018).
13. Peißl, S., Wickens, C. D., & Baruah, R. Eye-tracking measures in aviation: A selective literature review. *The International Journal of Aerospace Psychology*, **28**(3-4), 98-112 (2018).
14. Ziv, G. Gaze behavior and visual attention: A review of eye tracking studies in aviation. *The International Journal of Aviation Psychology*, ***26***(3-4), 75-104 (2016).

15. Ayala, N., Mardanbegi, D., Zafar, A., Niechwiej-Szwedo, E., Cao, S., Kearns, S., Irving, E & Duchowski, A. Does fiducial marker visibility impact task performance and information processing in novice and low-time pilots? *Computer and Graphics*, *199 (2024b)*.
16. Leube, A., & Rifai, K. Sampling rate influences saccade detection in mobile eye tracking of a reading task. *Journal of Eye Movement Research*, ***10***(3), (2017).
17. Taylor, R. M., & Selcon, S. J. Cognitive Quality and Situational Awareness with Advanced Aircraft Attitude Displays. *Proceedings of the Human Factors Society Annual Meeting*, ***34***(1), 26-30 (1990).
18. Recarte, M. Á., Pérez, E., Conchillo, Á., & Nunes, L. M. Mental workload and visual impairment: Differences between pupil, blink, and subjective rating. *The Spanish Journal of Psychology*, ***11***(2), 374-385 (2008).
19. Zheng, B., Jiang, X., Tien, G., Meneghetti, A., Panton, O. N. M., & Atkins, M. S. Workload assessment of surgeons: correlation between NASA TLX and blinks. *Surgical Endoscopy*, ***26***, 2746-2750 (2012).
20. Shannon, C.E.. A Mathematical Theory of Communication. In *The Bell System Technical Journal,* **27** (3), 379–423 (1984).
21. Shiferaw, B., Downey, L., & Crewther, D. A review of gaze entropy as a measure of visual scanning efficiency. *Neuroscience & Biobehavioral Reviews*, ***96***, 353-366 (2019).
22. Ciuperca, G., & Girardin, V. Estimation of the entropy rate of a countable Markov chain. *Communications in Statistics—Theory and Methods*, ***36***(14), 2543-2557 (2007).
23. Kowler, E. Eye movements: The past 25 years. *Journal of Vision Research, **51**,* (13), 1457-1483 (2011).
24. Diaz-Piedra, C., Rieiro, H., Cherino, A., Fuentes, L. J., Catena, A., & Di Stasi, L. L. The effects of flight complexity on gaze entropy: An experimental study with fighter pilots. Applied Ergonomics, **77**, 92-99 (2019).
25. Shiferaw, B. A., Downey, L. A., Westlake, J., Stevens, B., Rajaratnam, S. M., Berlowitz, D. J., ... & Howard, M. E. Stationary gaze entropy predicts lane departure events in sleep-deprived drivers. *Scientific Reports*, **8**(1), 1-10 (2018).
26. Parker, A. J., Kirkby, J. A., & Slattery, T. J. Undersweep fixations during reading in adults and children. *Journal of Experimental Child Psychology*, ***192***, 104788 (2020).
27. Federal Aviation Administration. Airplane Flying Handbook. *Washington, DC: Federal Aviation Administration* (FAA-H-8083-3C), (2021).
28. Brown, D. L., Vitense, H. S., Wetzel, P. A., & Anderson, G. M. Instrument scan strategies of F-117A Pilots. *Aviation, Space, and Environmental Medicine, **73***(10), 1007-1013 (2002).
29. Lu, T., Lou, Z., Shao, F., Li, Y., & You, X. Attention and Entropy in Simulated Flight with Varying Cognitive Loads. *Aerospace Medicine and Human Performance*, ***91***(6), 489-495 (2020).
30. Dehais, F., Peysakhovich, V., Scannella, S., Fongue, J., & Gateau, T. " Automation Surprise" in Aviation: Real-Time Solutions. In Proceedings of the 33rd annual ACM conference on Human Factors in Computing Systems (pp. 2525-2534), (2015, April).
31. Ayala, N., Keanrs, S., Irving, E., Cao, S., & Niechwiej-Szwedo, E. (submitted). The effects of a dual task on gaze behavior examined during a simulated flight in low-time pilots. *Frontiers in Psychology*.
32. Ayala, N., & Heath, M. Executive dysfunction after a sport-related concussion is independent of task-based symptom burden. *Journal of Neurotrauma*, **37**(23), 2558-2568 (2020).

33. Huddy, V. C., Hodgson, T. L., Kapasi, M., Mutsatsa, S. H., Harrison, I., Barnes, T. R., & Joyce, E. M. Gaze strategies during planning in first-episode psychosis. Journal of Abnormal Psychology, **116**(3), 589 (2007).
34. Irving, E. L., Steinbach, M. J., Lillakas, L., Babu, R. J., & Hutchings, N. (2006). Horizontal saccade dynamics across the human life span. *Investigative Ophthalmology & Visual Science,* **47**(6), 2478-2484 (2006).
35. Yep, R., Smorenburg, M. L., Riek, H. C., Calancie, O. G., Kirkpatrick, R. H., Perkins, J. E., ... & Munoz, D. P. Interleaved pro/anti-saccade behavior across the lifespan. *Frontiers in Aging Neuroscience*, 14, 842549 (2022).
36. Manoel, E. D. J., & Connolly, K. J. Variability and the development of skilled actions. *International Journal of Psychophysiology*, ***19***(2), 129-147 (1995).

NOTES FROM MEETING WITH EWA:

- JoVE
    - Ewa will write to editor and ask about purpose?
    - Next steps would be to gather data from all emergency scenarios, and re-write the manuscript
        - Include Allison's data and SIM 2 data
- WATS 2025
    - Data not there yet so wait until next year or the year after
- Motor Control
    - Summarize in an email to EWA what the top 5-6 things that we think are meaningful
    - What is useful to incorporate into their broader analyses- all strong analyses (especially throttle input and position)
- Gender paper
    - FINISH GOD DAMN
- ETRA
    - Either submit gender OR error trial data
    - Potential to look at entropy and k-coefficient for a later time