

Effects of Eye Measures on Human Controller Remnant and Control Behavior

Alexandru Popovici*
San Jose State University
NASA Ames Research Center
Delft University of Technology

Peter M. T. Zaal†
San Jose State University
NASA Ames Research Center

Daan M. Pool‡
Delft University of Technology

Max Mulder§
Delft University of Technology

Marinus M. van Paassen¶
Delft University of Technology

The aim of the current research was to investigate the possible relation between changes in eye activity parameters, variations in human remnant at the perceptual level and changes in human operator model parameters. Fourteen subjects performed a pitch tracking task, in which the display brightness was varied by changing the background color around a simplified primary flight display, in order to create a controlled, quasilinear change in the pupil diameter through the pupillary light reflex. Pupil diameter, blink, eye opening, and opening and closing amplitudes and speeds were recorded using an eye tracker. Participants controlled single integrator-like and double integrator-like dynamics. The variation in pupil diameter did not introduce significant differences in neither remnant characteristics nor the human operator model parameters. An interesting effect occurred in the human controller's time delay for the single integrator task, where the time delay was significantly higher for the darkest brightness compared to the other levels of brightness. This effect was not observed for the double integrator dynamics. Data suggested that the more difficult controlled dynamics induced a squinting effect, visible in smaller eye opening, and smaller eye opening and closing amplitudes. These results suggest that performance, and control behavior are invariant to the display brightness. Moreover, monitoring changes in the eye activity could represent a method of predicting variations in human remnant characteristics and human controller model parameters, introduced by task difficulty.

Nomenclature

A_f	forcing function amplitude, deg	H_p	human operator response
bl_{count}	blink count, -	H_{ol}	open-loop response
bl_{dur}	blink duration, ms	K_d	controlled dynamics gain, deg
cl_a	eye closing amplitude, mm	K_r	remnant gain, -
cl_s	eye closing speed, cm/s	K_v	visual gain, 1/deg
e	error signal, deg	n_f	forcing function prime, -
eye_{op}	eye opening, mm	op_a	eye opening amplitude, mm
f_f	forcing function, deg	op_s	eye opening speed, cm/s
H_c	controlled dynamics response	pup_d	pupil diameter, mm

*Research Scholar, Human Systems Integration Division, NASA Ames Research Center, Moffett Field, CA, 94035; alexandru.popovici@nasa.gov

†Senior Research Engineer, Human Systems Integration Division, NASA Ames Research Center, Moffett Field, CA, 94035; peter.m.t.zaal@nasa.gov. Member.

‡Associate Professor, Faculty of Aerospace Engineering, Delft University of Technology, Delft, The Netherlands; d.m.pool@tudelft.nl. Member.

§Professor, Faculty of Aerospace Engineering, Delft University of Technology, Delft, The Netherlands; m.mulder@tudelft.nl. Associate fellow.

¶Assistant Professor, Faculty of Aerospace Engineering, Delft University of Technology, Delft, The Netherlands; m.m.vanpaassen@tudelft.nl. Member.

R_e	remnant injected at e , deg	<i>Symbols</i>	
R_u	remnant injected at u , deg	ζ_n	neuromuscular damping, -
s	Laplace operator	τ_v	visual time delay, ms
S_{ff}	power spectrum of f_f , deg ² s/rad	ϕ_f	forcing function phase, deg
S_{uu}	power spectrum of u , deg ² s/rad	ϕ_m	open-loop phase margin, deg
$S_{uu,f}$	S_{uu} due to f_f , deg ² s/rad	$\Phi_{rr,e}$	remnant power spectrum, deg ² s/rad
$S_{uu,r}$	S_{uu} due to R_e , deg ² s/rad	$\dot{\Phi}_{rr,e}$	normalized $\Phi_{rr,e}$, s/rad
T_d	controlled dynamics time constant, s	θ	output pitch angle, deg
T_L	visual lead time constant, s	ω	frequency, rad/s
T_m	measurement time, s	ω_c	open-loop crossover frequency, rad/s
T_r	remnant time constant, s	ω_i	forcing function bandwidth, rad/s
u	control input signal, -	ω_f	forcing function frequency, rad/s
		ω_m	measurement time base frequency, rad/s
		ω_n	neuromuscular frequency, rad/s
		ω_r	remnant break frequency, rad/s

I. Introduction

Linear transfer functions have been widely used to model human manual control behavior and can explain a large part of the mechanism behind it.¹ The part that cannot be modeled by linear transfer functions, the remnant, can be attributed to different sources, related to system noise and the exploratory nature of human behavior. A few examples include true observation noise (the error in observing task variables), motor noise, nonlinearities in the human controller (time-varying parameters, time delays), or aperiodic sampling of the perceived variables.² As Flach mentions in his work on active psychophysics, the part of the control signal linearly correlated to the input gives information on the task and performance of the human operator, whereas the remnant gives an insight into the human operator himself.³ Levison et al. concluded that the remnant can be represented as an equivalent observation noise injected at the human operator's perceptual level.² They hypothesized that the remnant is in fact white noise injected on each of the perceived variables. The variation of remnant characteristics with most task variables is relatively well studied and understood.⁴ However, no previous studies looked at the possible relation between physiological eye parameters and remnant characteristics. The eyes being the channel through which visual perception is achieved, it is interesting to investigate if changes in the eye parameters create changes in the human perceptual remnant.

The importance of a better understanding of the human remnant is three-fold. First, identifying and explaining certain sources of the remnant will give a better insight on whether the remnant is of physiological nature, an intrinsic perceptual process or a combination of both. Second, more complex human operator models that account for human variability and other external environmental factors can be developed in order to better predict human manual control behavior. Lastly, correctly identifying remnant sources could potentially be used in predicting human operator workload and state.

The goal of this paper is to investigate the possible relation between changes in physiological eye parameters, changes in remnant model characteristics and human operator model parameters. The approach consists of a pitch manual tracking task in which an eye tracker was used to capture changes in eye activity parameters, such as blink count and duration, pupil diameter, eye opening and closing amplitudes and speeds. Pupil diameter was controlled through the activation of the pupillary light reflex, by varying the color of the background around a simplified primary flight display (PFD). Furthermore, two different controlled dynamics were considered: simple, single integrator-like dynamics, and more difficult, double integrator-like dynamics, to investigate if the task difficulty has an impact on the investigated relations. Lastly, the human remnant was modeled according to Levison's representation, as an equivalent observation noise injected into the error signal. This model resembles a squared low pass filter in the frequency domain, being composed of a gain and a break frequency.

The paper is structured as follows. First, an overview of the pitch tracking task is given, including an explanation of different task variables and the remnant model used. Next, a method section describes the experiment design and data analysis. At the end of this section, several hypotheses are formulated. Following, the results of the experiment are presented. These findings are covered in detail in the discussion section. The paper ends with conclusions and references.

II. Control task

In this experiment, a compensatory pitch tracking task was considered, in which an eye tracker was used to collect eye measurements. The control diagram is depicted in Fig. 1. The goal of the human operator was to minimize the error e presented on a simplified primary flight display (PFD), by providing control inputs u with a joystick. These inputs were transformed into the output θ through the controlled dynamics transfer function H_c . The error e represents the difference between the target forcing function f_f and the output θ .

The quasi-linear human operator model is typically composed of two parts.¹ A linear transfer function, H_p is used to model equalization dynamics, time delays, and neuromuscular limitations. The second part is the remnant signal, R_e which accounts for human controller's nonlinear behavior. It represents a signal in the control loop that embodies control behavior that is not linearly correlated to the input forcing function. The remnant can be represented anywhere in the control loop, for example also at the control input u , as marked by R_u , due to the linear nature of the other transfer functions. However, representing the remnant at the perceptual level of the human operator through R_e allows for a straightforward comparison of remnant characteristics with changes in eye parameters. Note that, regardless of where it is injected in the control loop, the remnant signal represents a lumped contribution of nonlinearities from all sources, as it is impossible to distinguish between them.

In the remainder of this section, the different elements in the control loop in Fig. 1 are described.

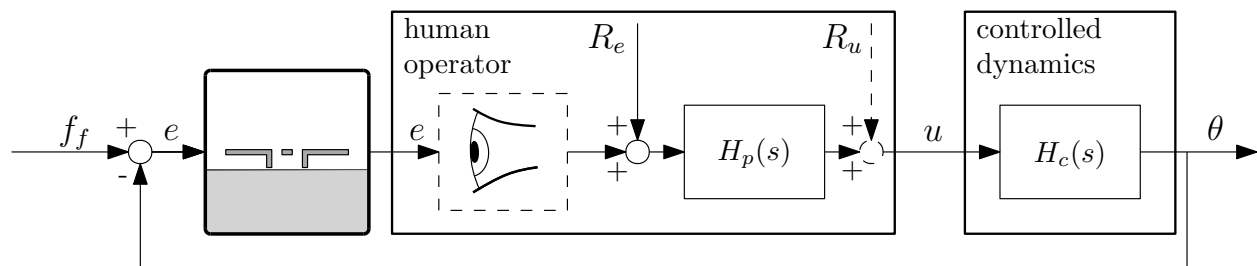


Figure 1. Compensatory tracking task control diagram.

1. Controlled dynamics

Two different controlled dynamics were considered in this experiment. They were varied between a simple, single integrator-like and a more difficult double integrator-like dynamics. The transfer function of both controlled dynamics is given by Eq. (1).

$$H_c(s) = \frac{K_d}{s(sT_d + 1)}, \quad (1)$$

where K_d is a gain and T_d a time constant that indicates at which frequency ($\omega_d = 1/T_d$) the controlled dynamics become a double integrator. The choice for this model for the controlled dynamics will later be explained in Section 3. The magnitude and phase of their frequency response functions are shown in Fig. 2a and Fig. 2b.

The vertical, thinner lines indicate the location of the dynamics break frequency ω_d . The response indicated with "SI-like" dynamics is easier to control, with a higher break frequency of 6 rad/s, whereas for the "DI-like" dynamics the break frequency is 1 rad/s. The smaller the break frequency of the controlled dynamics, the harder the task becomes since frequencies in the control input are suppressed earlier, not contributing to the output θ . Therefore, visual lead has to be generated at earlier frequencies.¹ Also note the increased phase lag for the DI-like dynamics. The gain of the dynamics transfer functions was chosen such that the control authority was similar for both controlled dynamics.

2. Human operator transfer function

Most of human control behavior in this type of tracking tasks can be captured by a linear transfer function that captures the visual response of the human controller. Moreover, according to the crossover model, humans adjust their control strategy such that the human controller-vehicle dynamics open-loop frequency

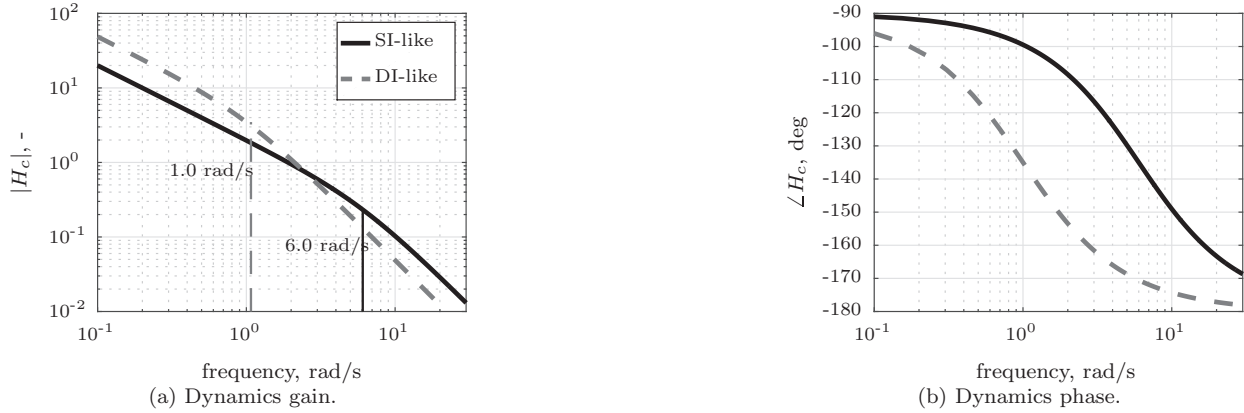


Figure 2. Controlled dynamics frequency response functions.

response function approximates single integrator dynamics near the crossover frequency.¹ Considering the controlled dynamics presented in Section 1, visual lead is required near the break frequencies of the controlled dynamics.¹ The human operator transfer function then becomes:

$$H_p(s) = K_v(1 + T_L s)e^{-\tau_v s} \frac{\omega_n^2}{\omega_n^2 + 2\zeta_n \omega_n s + s^2} \quad (2)$$

Visual gain K_v and visual lead time constant T_L parametrize the human operator's equalization used to achieve the desired performance, whereas physical limitations are characterized by the visual time delay τ_v , neuromuscular damping ζ_n and neuromuscular frequency ω_n .

Measures of stability and performance can be obtained by analyzing the crossover frequency and phase margin of the human-vehicle open-loop frequency response function, indicated in Eq. (3).

$$H_{ol}(j\omega) = H_p(j\omega)H_c(j\omega) \quad (3)$$

3. Remnant

To investigate if changes in eye activity affect human remnant, a model to capture remnant characteristics is needed. In their early work, McRuer et al. concluded that the remnant power spectral density is a smooth function of frequency and that its most consistent representation is as an equivalent observation noise injected at the controller's input.⁵ Moreover, their work showed that the order of the controlled dynamics heavily impacts the remnant spectrum. In addition, Pew et. al reported that the remnant spectrum is invariant to the bandwidth of the input forcing functions and display gain.⁴ Later research confirmed these findings, showing that an equivalent observation noise at the human controller's input, normalized by the variance of the error, is invariant to control input characteristics, ultimately meaning that the absolute remnant power scales with the magnitude of the error.⁶

Levison et al. concluded that, if noise signals following Weber's law act on each of the perceived variables in a tracking task (typically error position e and rate \dot{e}), then perceptual remnant can be represented as an equivalent observation noise injected on the error signal e . Then, this equivalent remnant spectrum at the perceptual level, normalized by the variance of the error position σ_e^2 , can be represented by a squared first-order low pass filter model. The power spectrum of such model is given by:

$$\dot{\Phi}_{rr,e}(\omega) = \frac{\Phi_{rr,e}^2}{\sigma_e^2} = \frac{K_r}{1 + T_r^2 \omega^2} = \frac{K_r}{1 + \frac{1}{\omega_r^2} \omega^2} \quad (4)$$

where K_r represents the gain of the remnant, T_r is a constant that dictates the ratio in the perception gains on the error rate (\dot{e}) and error displacement (e), and ω denotes frequency. Note that the remnant filter break frequency is then given by $\omega_r = 1/T_r$, and is expected to change with the controlled dynamics, due to the variation of the perception gains on \dot{e} and e introduced by the equalization of the human operator. The variation of K_r with the controlled dynamics is more difficult to predict, since it not only depends on

the perception gains on \dot{e} and e , but also on the variance of the error position and rate signals.² In Eq. (4), $\Phi_{rr,e}$ represents the power spectral density of the signal R_e depicted in Fig. 1.

This model has been validated by Jex et. al, in a study that compiled remnant data obtained from numerous experiments.⁷ As mentioned before, the remnant characteristics are affected by controlled dynamics.⁵ It was experimentally found that, for controlled dynamics of the form $H_c(s) = K_d/s(sT_d + 1)$, where the time constant T_d is neither too small or too large, the remnant model has two identifiable parameters, K_r and ω_r . However, in case T_d approaches zero or infinity, the remnant break frequency will also become infinity or zero, respectively, meaning that the only identifiable parameter in Eq. (4) is the gain K_r . This is the result of the ratio between the weightings applied by the human controller on perceiving the error rate and error position, which is affected by the order of the controlled dynamics. In the present experiment, T_d was chosen such that both the remnant gain and break frequency can be identified from the power spectrum of the remnant, for both the single integrator-like and double integrator-like controlled dynamics.

4. Visual display

In order to create a controlled change at the visual perceptual level, different levels of display brightness were used in order to activate the pupillary light reflex which dilates and constricts the pupil, regulating the amount of light entering the eye.⁸ Pupil size is the eye parameter of choice since activating the pupillary light reflex represents a non-intrusive and straightforward way to create a controlled change of at least one parameter at the perceptual level.

Different levels of luminance (perceived brightness) can be achieved by varying the background color of the display on which the simplified PFD is represented. Fig. 3 shows the layout of the display presented.

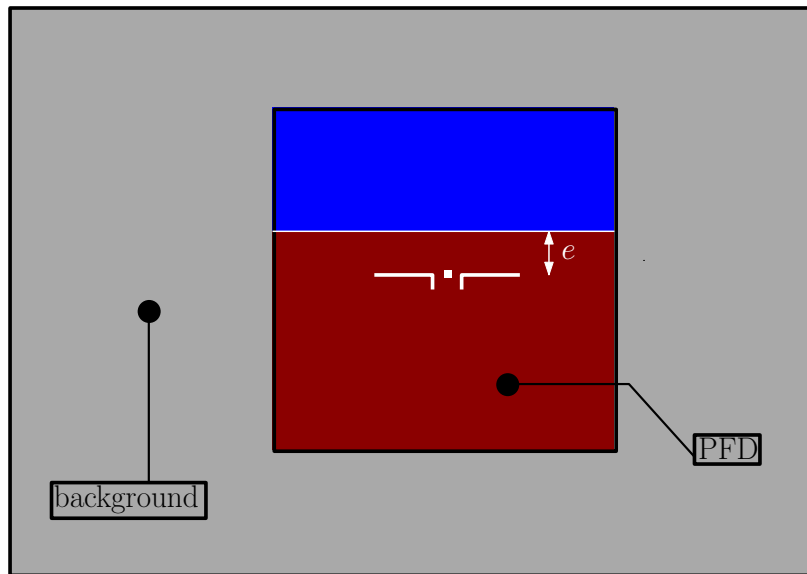


Figure 3. Primary Flight Display, with the background color around it.

In the figure above, the display containing a simplified PFD and the background color around it are shown. The error signal e is also depicted in the figure. In this case, the background color is light grey. The total luminance of this display can be changed by altering the background color. It was preferred to choose several greyscale RGB values for the background color (thus, brightness) that result in a linear change in the pupil diameter, in order to achieve the highest variation between experimental conditions. Pupil diameter is not a linear function of luminance, however, as shown by several studies that attempted to create a mathematical model of pupil variation with light intensity.⁹ Furthermore, the luminance perceived is composed of the light coming from the display background color and the PFD itself. Lastly, the perceived brightness also depends on where the subject is looking at a given time. Therefore, in order to achieve a linear change across several brightness levels, the RGB values were chosen empirically, analyzing the pupil response from the first author of this study in a small experiment. The procedure followed is summarized below.

- The greyscale RGB values $[v, v, v]$ of the display background color were changed from $[0, 0, 0]$ - black background to $[1, 1, 1]$ - white background, in steps of $[0.33 \ 0.33 \ 0.33]$
- A tracking task was performed for at least 30 seconds for each RGB set, in order to allow for the pupil diameter to achieve a stable level
- Pupil diameter was recorded for all the RGB levels presented above, and an exponential model of the form given by Eq. (5) was fit through the measurements, normalized between 0 and 1, where a , b and c are parameters of the exponential curve, v the RGB value and p the pupil diameter
- A number of brightness levels ($N = 4$) was chosen, and RGB values (v_i, v_i, v_i) that would result in a linear change in pupil diameter. They were calculated using Eq. (6), with the starting RGB value, v_1 , equal to $[0.0, 0.0, 0.0]$.

$$p(v) = a + be^{-cv} \quad (5)$$

$$v_i = -\frac{1}{c} \ln\left(1 - \frac{i}{N-1} \frac{1}{b}\right), \quad i = 2, \dots, N \quad (6)$$

Fig. 4 shows an example of the obtained RGB values ($[v_1, v_1, v_1]$, $[v_2, v_2, v_2]$, $[v_3, v_3, v_3]$ and $[v_4, v_4, v_4]$) that would potentially result in a linear change in pupil diameter. Analyzing the figure, it can be seen that $p(v_1) - p(v_2) = p(v_2) - p(v_3) = p(v_3) - p(v_4)$. For this specific example, $N=4$ (four different pupil diameters, linearly spaced), and the resulting parameters from the model fit were $a=-0.17$, $b=1.18$, and $c=2$. Using Eq. (6), the obtained RGB levels were $v_1=0.0$, $v_2=0.17$, $v_3=0.42$, and $v_4=0.94$.

The described method does not guarantee a linear change in pupil diameter, since the pupillary light reflex might be different from individual to individual. However, it provides an initial estimation for the background colors that would result in clear differences in pupil diameter.

5. Eye measures

The eye tracker recorded several eye parameters during the experiment: pupil diameter, blink count and duration, eye opening, opening and closing amplitude and speeds. A detailed description on how these parameters are obtained can be found in Appendix A. As mentioned before, Levison hypothesized that in a tracking task, one possible source for the remnant is in fact represented by white noise injected on each of the perceived variables.² By comparing changes in the recorded eye parameters with the remnant behavior, the possible relation between the two can be investigated. It could be that the amount of light entering the eye (inferred from the pupil diameter) affects the power of this white noise, which in turn affects the remnant gain. Another possibility is that an increase in blink count and/or duration could affect the ratio between the perceptual gains on error velocity and position, resulting in changes in the break frequency of the remnant. Since no previous studies looked at such relations, the evolution of all the eye measures mentioned above will be recorded from the eye tracker.

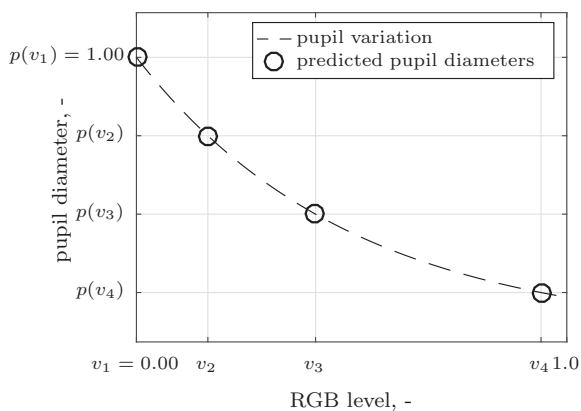


Figure 4. Predicted RGB values for a linear change in pupil diameter.

III. Method

A. Experiment design

1. Experimental conditions

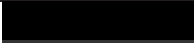





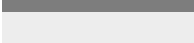
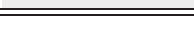
The two independent variables manipulated in the experiment were the controlled dynamics (DYN), and display brightness (BR). Two different controlled dynamics and four different levels of brightness were tested. Table 1 provides an overview of all the experimental conditions.

In the first four conditions (C1-C4), the task was to control single integrator-like dynamics (SI), while the display brightness was varied across four levels (BR1-BR4). The same brightness settings were tested with more difficult, double integrator-like dynamics (DI).

The main difference in the controlled dynamics is given by ω_d , the frequency at which the frequency response of the dynamics transits from a single-integrator to a double integrator (Fig. 2). For the task that was easier to control, ω_d was higher, meaning that the dynamics were single integrator-like for a wider frequency range. For the more difficult task, ω_d was much smaller. As mentioned previously, the gain K_d was selected such that the control authority was similar for both controlled dynamics.

The display brightness was varied by changing the grayscale background color of the PFD. The RGB values, together with the resulting background color used to create a quasi-linear change in pupil diameter are provided in Table 1. Background luminance was measured using a Konica Minolta LS-100 luminance meter. It is important to mention that, in order to keep the same level of luminance regardless of the magnitude of the pitch angle displayed in Fig. 3, the colors of the PFD sky and ground planes were isoluminant, with a luminance of 6.5 cd/m^2 .

Table 1. Experimental conditions.

Condition	DYN	BR	K_d, ω_d [rad, rad/s]	Transfer function	(R, G, B)	Luminance [cd/m ²]	Background Color		
C1		BR1			(0.00, 0.00, 0.00)	0.07			
C2	SI	BR2	(0.8, 6.0)	$\frac{0.8}{s(s/6+1)}$	(0.20, 0.20, 0.20)	2.60			
C3		BR3					(0.49, 0.49, 0.49)	18.30	
C4		BR4					(0.93, 0.93, 0.93)	60.80	
C5		BR1					(0.00, 0.00, 0.00)	0.07	
C6	DI	BR2	(0.196, 1.0)	$\frac{0.196}{s(s+1)}$	(0.20, 0.20, 0.20)	2.60			
C7		BR3					(0.49, 0.49, 0.49)	18.30	
C8		BR4					(0.93, 0.93, 0.93)	60.80	

2. Participants

Fourteen general aviation pilots participated in the experiment. Two pilots had significantly more flying hours than others, with 1200 and 2000 hours, respectively. The average number of flight hours was 366 hours with a standard deviation of ± 585 hours. Participant age was between 19 and 35 years, with the mean age of 27 years and a standard deviation of ± 5 years. One participant did not have the general aviation license, however he/she had more than five hours tracking tasks experience in fixed-based simulators.

3. Apparatus

Two different systems were used for the experiment, one running the tracking task, and one running the eye tracker software. The tracking task was performed on a Microsoft Windows system running at 100 Hz. The control loop was developed using the control systems library for Python. For visualization, the OpenGL library was used. The computer was equipped with a LCD display having a resolution of 1920x1200 pixels, set at the lowest brightness. The PFD shown had a physical size of 15x15 cm. A BG Systems Inc. JFx joystick was used for making control inputs. The joystick was unlocked in all axes. However, only the pitch commands were used in the task. Joystick output was normalized to the -1 and 1 range, corresponding to full pitch up and pitch down inputs, respectively.

For eye tracking, a SmartEye Pro system collecting data at 60 Hz was used on a different desktop computer. The subjects were illuminated with two sets of infrared LED arrays. Four different infrared cameras that captured the head's position and rotation in the physical space were used for the eye tracking. The tracking task and the eye tracker systems communicated via an UDP connection, using a network switch.

An overview of the experimental setup is depicted in Fig. 5, where both the pitch tracking task and the eye tracker setup are visible. In addition, a more detailed description of the applications used for the experiment and their interaction is given in Appendix B.

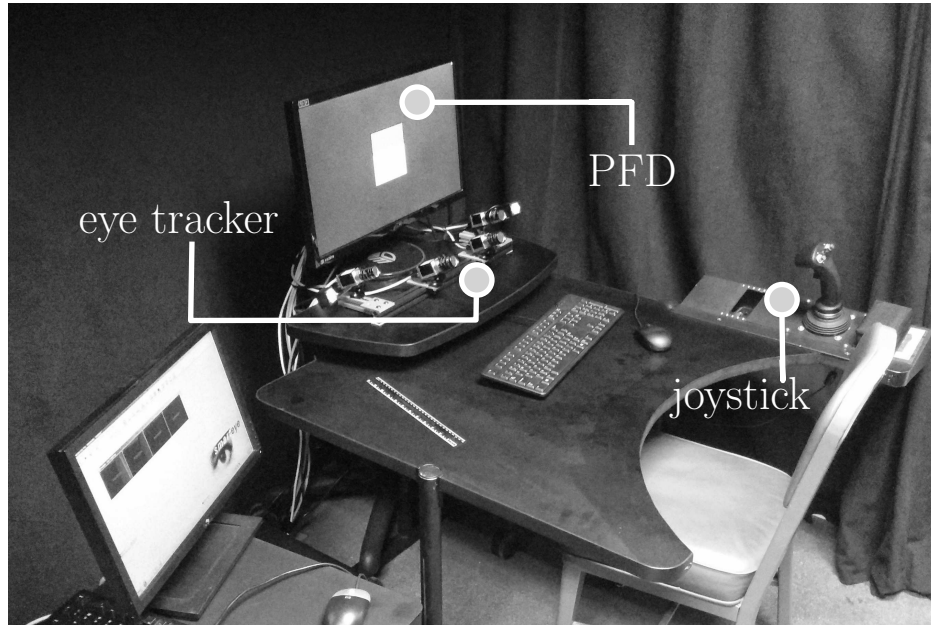


Figure 5. Experiment setup.

4. Procedure

At the beginning of the experiment, participants received a thorough briefing, being instructed to follow the forcing function as closely as possible, by minimizing the tracking error, keeping the aircraft symbol on the PFD aligned with the horizon (Fig. 3). Examples of bad and good control inputs were demonstrated prior to the beginning of the experiment. Bad control inputs were indicated as "bang-bang" inputs, no inputs at all, or inputs in which the strategy changed drastically during individual runs. Good inputs were indicated as smooth, high-frequency content inputs. Pilots were encouraged to obtain a better score than the highest score obtained by previous participants. Financial compensation was provided to all participants.

Calibration of the eye tracker was performed next. Participants were instructed to sit in the most comfortable position for the experiment, while making sure that their facial features were visible in all four infrared cameras. The last step consisted of participants slightly moving and rotating their heads while keeping their focus in the center of the display. This allowed the eye tracker software to create a model of the participant's face and head that is robust to small head movements and rotations, during the experiment. This procedure ensured a higher accuracy in the collection of the eye measurements.

All participants performed 80 runs, each lasting 90 seconds. The first 40 runs were either a latin square distribution of conditions C1 to C4, or C5 to C8. In the last 40 runs pilots performed the other controlled dynamics, with a latin square distribution of the four possible brightnesses. That is, in order to obtain a balanced experiment, eliminating the potential effects on control behavior of performing certain controlled dynamics first, half of the subjects did conditions C1-C4 first, and the other half C5-C8 first. This distribution resulted in participants seeing each brightness ten times for each controlled dynamics in the complete experiment. The first 16 runs for each dynamics were used as training. In the data analysis, only the last 24 runs for each controlled dynamics (six runs for each brightness) were used.

After each run, the display background color changed, and a 15-second break was indicated as a count-

down, in order to allow participants' pupil diameter to adapt to the new brightness. Moreover, subjects could take a three to five minutes break after each set of eight runs, and a 15 minutes break after 40 runs. The score, calculated as the root mean square (RMS) of the tracking error, was displayed on the PFD after each run. Participants were encouraged to obtain a value as small as possible for the score, indicating a low RMS of the tracking error, and thus a good performance.

At the end of the experiment, subjects were asked to rate the level of discomfort caused by each display brightness ("annoyance factor") on an ordinal, Likert-like scale from 1 to 4. A value of 1 indicated the brightness that resulted in the smallest discomfort while controlling the task, whereas a value of 4 was a measure of the most irritating level of brightness. The rating considered the SI and DI tasks together.

5. Forcing function

The forcing function f_f was composed of a summation of ten sine-waves with different amplitudes, frequencies and phases, as indicated in Eq. (7):

$$f_f(t) = \sum_{k=1}^{10} A_f(k) \sin[\omega_f(k)t + \phi_f(k)], \quad (7)$$

where $A_f(k)$, $\omega_f(k)$, and $\phi_f(k)$ represent the amplitude, frequency and phase of individual sine-waves, and t the time variable. Table 2 gives an overview of the ten components of the forcing function.

Table 2. Overview of the forcing function properties.

$n_f, -$	$\omega_f, \text{rad/s}$	A_f, deg	ϕ_f, deg
6	0.460	6.2472	-84.774
13	0.997	4.3688	-4.269
27	2.070	1.9712	40.141
41	3.144	1.0616	-112.088
53	4.065	0.7128	-161.179
73	5.599	0.4416	120.470
103	7.900	0.2808	-149.989
139	10.661	0.2048	129.202
174	13.346	0.1712	-38.612
229	17.564	0.1456	11.127

The measurement time in each run used for data analysis was $T_m = 81.92$ seconds, resulting in a base frequency of $\omega_m = 2\pi/T_m = 0.0767$ rad/s. The frequencies of the individual sine-waves, $\omega_f(k)$, were integer multiples n_f of the base frequency ω_m . The amplitudes were determined using a second-order low-pass filter that reduces the amplitudes of the forcing function at high frequencies. This results in a task that is not extremely difficult to control, even when controlling double integrator-like dynamics. The integer multiples, together with the second order low-pass filter used for determining the amplitudes were successfully used to identify human control behavior in another experiment.¹⁰ Furthermore, the phases chosen resulted in the forcing function having an average crest factor.¹¹

6. Dependent measures

One subjective measure indicating the level of discomfort caused by different levels of brightness was analyzed ("annoyance factor"). Thirteen subjective measures were the remnant model parameters and total power ($K_r, \omega_r, \sigma_{\Phi_{rr,e}}^2$), RMS of tracking error and control input (RMS_e, RMS_u), human operator model variables ($K_v, T_L, \tau_v, \zeta_n, \omega_n$), the variance accounted for (VAF) of the control input, and performance and stability measures (ω_c, ϕ_m). Finally, the eye tracker recorded eight eye parameters during the experiment: pupil diameter (pup_d), eyelid opening (eye_{op}), blink count and duration (bl_{count}, bl_{dur}), and eye opening and closing amplitudes and speeds (op_a, op_s, cl_a, cl_s). An explanation on the nature of the collected eye parameters is given in Appendix A.

7. Hypotheses

Several hypotheses were formulated, based on previous research. First, increasing the level of display brightness should result in an almost linear decrease in pupil diameter caused by the pupillary light reflex (H1). Furthermore, it was hypothesized that the changes in pupil diameter will be invariant to the difficulty of the controlled dynamics, due to the fact that a manual tracking task involves skill-based behavior, the more unstable dynamics not requiring more cognitive workload that would potentially increase the pupil diameter¹² (H2). However, more continuous attention would be required to perceive error position and rate, which will

affect the gains applied on these variables. Total remnant power was expected to be larger in the DI task, compared to the SI task. In addition, the DI task would result in a lower remnant break frequency and a higher remnant gain (H3).⁷ It is also hypothesized that the potential effects of changes in eye parameters on the remnant characteristics would be rather small, and remnant characteristics would be affected by some of the eye activity measures only (H4). Finally, subjective ratings on how much discomfort is caused by a certain display background would have an effect on the human remnant and performance. It is hypothesized that the most irritating background colors will result in an increase in both nonlinear behavior and tracking error (H5).

B. Data Processing

1. Data acquisition

To obtain the dependent measures described in Section III.A.6, several methods were used, as summarized in this subsection. The last six out of ten recorded runs, for each brightness and dynamics, were used to calculate all the dependent measures.

HUMAN OPERATOR MODEL PARAMETERS For estimating the human operator model parameters defined in Eq. (2), the time domain data of the error signal e and control input u were averaged over six runs for each condition, in order to minimize the contribution of the stochastic remnant signal, that increases the bias in parameter estimation. A time-domain parameter estimation technique that uses maximum likelihood estimation was applied in order to obtain the parameter vector $[K_v T_L \tau_v \zeta_n \omega_n]$.¹³ This method uses a genetic algorithm in order to determine initial parameter estimates, which are then refined through a gradient-based Gauss-Newton estimation. A set of human controller model parameters is obtained for each of the conditions summarized in Table 1. Once a parameter set is found, the crossover frequency and phase margin can be obtained from the human operator-controlled dynamics open-loop transfer function.

REMNANT AT THE HUMAN PERCEPTUAL LEVEL The remnant at the perceptual level cannot be directly measured in the experiment, since it is a process internal to the human operator. However, it can be retrieved from signals circulating in the loop.² Due to the linear nature of the transfer functions, it is then possible to obtain the remnant signal at any position in the control loop. In this case, the location of the tracking error e is chosen, as previously indicated in Fig. 1.

Eq. (8) - Eq. (13) show how to obtain the remnant spectrum at the forcing function frequencies ω_f , from the control input u :

$$S_{uu}(\omega_f) = S_{uu,f}(\omega_f) + S_{uu,r}(\omega_f) \quad (8)$$

$$S_{uu,r}(\omega_f) = \sum_{\substack{\omega=\omega_f-2 \\ \omega \neq \omega_f}}^{\omega_f+2} S_{uu}(\omega)/4 = \quad (9)$$

$$= \Phi_{rr,e}(\omega_f) \frac{|H_p(\omega_f)|^2}{|1 + H_p(\omega_f)H_c(\omega_f)|^2} \quad (10)$$

$$S_{uu,f}(\omega_f) = S_{ff}(\omega_f) \frac{|H_p(\omega_f)|^2}{|1 + H_p(\omega_f)H_c(\omega_f)|^2} \quad (11)$$

$$\Rightarrow \Phi_{rr,e}(\omega_f) = S_{ff}(\omega_f) \frac{S_{uu,r}(\omega_f)}{S_{uu,f}(\omega_f)} \quad (12)$$

$$\dot{\Phi}_{rr,e}(\omega_f) = \Phi_{rr,e}(\omega_f)/\sigma_e^2 \quad (13)$$

In Eq. (8), $S_{uu}(\omega_f)$ represents the power spectral density of the control signal u at the frequencies ω_f of the forcing function. At these frequencies, the total power of the control signal u is composed of the power due to the forcing function $S_{uu,f}(\omega_f)$, and the remnant contribution $S_{uu,r}(\omega_f)$. Since the remnant is a continuous function of frequency,⁵ its power at the location of u at the input frequencies can be obtained as the average power of the control input at the neighboring frequencies, shown in Eq. (9).¹⁴ Furthermore, the remnant power at the control input can be related to the remnant at the perceptual level, $\Phi_{rr,e}(\omega_f)$, through Eq. (10), where $H_c(\omega_f)$ are the controlled dynamics and $H_p(\omega_f)$ the human operator transfer function.

Similarly, the forcing function power in the control input signal can be related to the power spectral density of the forcing function itself, as shown in Eq. (11). From Eq. (10) and Eq. (11), the remnant spectrum at the perceptual level can then be obtained directly from the power spectral densities of the input forcing function and the control input (Eq. (12)). Lastly, the normalized remnant at the perceptual level is obtained by dividing it with the variance of the control signal, σ_e^2 , shown in Eq. (13).

Normalized remnant power spectra ($\widehat{\Phi}_{rr,e}(\omega_f)$) are obtained for each of the six analyzed runs, and then averaged in the frequency domain, in order to obtain a characteristic remnant power spectrum for each combination of brightness and controlled dynamics tested in the experiment.

Once the remnant spectrum is obtained, parameters K_r and ω_r from Eq. (4) have to be obtained from the frequency domain data. A model is fit using an interior-point algorithm, in which the cost function takes into account the variability of the data at each frequency. Since the power spectrum $\widehat{\Phi}_{rr,e}(\omega_f)$ is obtained by averaging data from six runs in the frequency domain, the frequencies at which the variability is high can be given less weight in the cost function. The method of weighing the cost function with the variance of the power spectra at a certain frequency was successfully used in the past to reduce uncertainty in the model fit.¹⁵ The cost function that quantifies the difference between the measured power spectrum and the obtained model power spectrum is given in Eq. (14):

$$J(\theta) = \sum_{i=1}^{10} \frac{|\widehat{\Phi}_{rr,e}(\omega_f) - \widetilde{\Phi}_{rr,e}(\omega_f; \theta)|^2}{\sigma_{\widehat{\Phi}_{rr,e}(\omega_f)}^2} \quad (14)$$

When there is a large uncertainty at a certain frequency given by a large variance ($\sigma_{\widehat{\Phi}_{rr,e}}^2$), the error between the obtained model ($\widetilde{\Phi}_{rr,e}$) and the measured data in the frequency domain ($\widehat{\Phi}_{rr,e}$) will not heavily contribute to the cost function. In Eq. (14), $\theta = [K_r \ \omega_r]$ represents the parameter vector for the model fit. Fig. 6 shows an example of a remnant model fit, taken into account the variability of the remnant power spectra at each frequency. The light grey curves indicate the individual remnants, the black circles represent the mean remnant, the error bars the 95% confidence intervals of the mean at each frequency, and the dashed black line the model fit. Note that the error bars are not symmetrical in the plot due to the logarithmic axes.

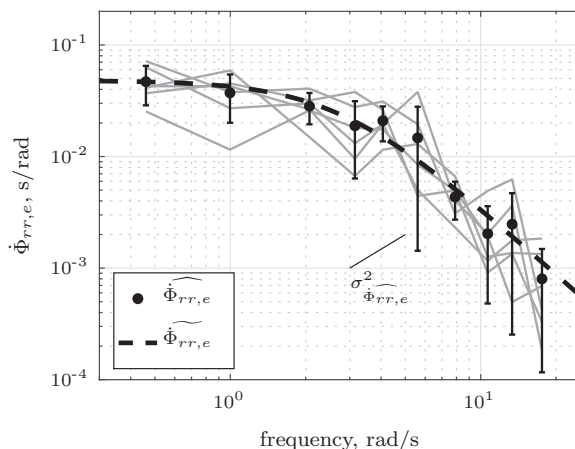


Figure 6. Example of a remnant spectrum model fit.

MEAN EYE ACTIVITY PARAMETERS The last 81.92 seconds in the same six runs were used to obtain the data for the eye parameters. Mean values for each experimental condition were obtained by averaging the time domain data of the six runs. That means, a single representative value was obtained for each eye measure, for each condition and subject. As an example, mean blink duration for subject 3 for condition C1 was 373 ms.

2. Statistical analysis

A two-way repeated measures Analysis of Variance (ANOVA) was performed to determine if any significant differences were introduced by the display luminance and controlled dynamics on all the dependent measures mentioned in Section 6. The two within-subject factors are represented by the controlled dynamics (DYN) and the brightness level (BR). The controlled dynamics had two levels, whereas brightness had four different levels, as summarized in Table 1 from Section III A.

Outliers that resulted from values greater than ± 3 standard deviations were removed. There were few outliers in the remnant gain and break frequency parameters, which were kept in the analysis, since they did not result from erroneous model fits. The majority of the dependent measures were normally distributed as assessed by Shapiro-Wilk's test of normality on the studentized residuals. Sphericity was tested using Mauchly's test of sphericity. In case the sphericity assumption was violated ($p < 0.05$), the Greenhouse-Geisser correction was applied.

First, it was determined whether there was a two-way interaction between dynamics and brightness. If a significant interaction existed, simple main effects were investigated by performing a one-way repeated measures ANOVA for each of the within-subject factors. If no significant interaction was found, significant main effects for the two within-subjects factors were interpreted.

Data for all the dependent measures were collected from 14 participants. For one participant, the data for the remnant parameters was not used for double integrator-like dynamics, since the remnant spectrum shape could not be fit with the proposed model. Moreover, one participant blinked very rarely throughout the experiment, which resulted in missing data for parameters affected by blink: blink duration, opening and closing amplitudes and speeds. These data were not used in the analysis, and considered as missing cases in the ANOVA analysis, where the degrees of freedom are reduced for the respective dependent measures.

IV. Results

When applicable, in all figures in this section markers represent the means and the error bars the 95% confidence intervals of the mean for each dependent measure. The four different levels of brightness are shown on the horizontal axis (BR1 - BR4), whereas the two different dynamics are indicated with two different markers (SI and DI, respectively). Between-subject variability was removed before calculating the confidence intervals.

Table 3 in Appendix C provides a summary of the two-way repeated measures ANOVA on all the dependent measures.

A. Subjective ratings

Participants were asked to rank the discomfort of each brightness setting, on a scale from 1 to 4. The results are shown in Fig. 7, where the sums of subjective ratings from all subjects for each brightness are presented. In general, participants found BR2 and BR3 the most comfortable, and the extreme cases in BR1 and BR4 the least comfortable. Brightness level 2 was the overall preferred brightness level, with the lowest level of discomfort, and brightness level 4, with a white background, was found to cause the most discomfort.

B. Performance and control activity

Fig. 8a shows the RMS value of the error. The overall RMS of the control error e is higher on average by about 3.5 degrees for the DI condition due to the more unstable dynamics, which made the task more difficult to control. Tracking error does not seem to be influenced by the different levels in brightness. There was no statistically significant two-way interaction for the error RMS, as expected by inspecting Fig. 8a. The main effects of the controlled dynamics were statistically significant, $F(1, 13) = 216.00, p < 0.05$. No significant difference was found for the main effect of the brightness level.

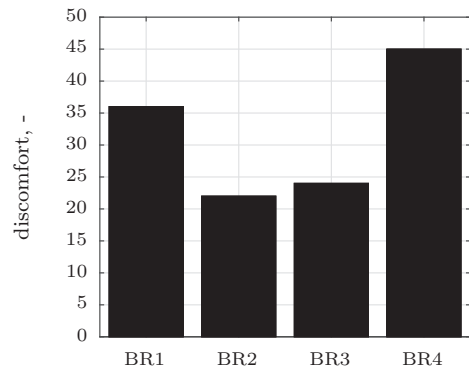


Figure 7. Subjective ratings for each brightness level.

The RMS of the control input is shown in Fig. 8b, where no notable trends can be observed. No significant two-way interaction between the dynamics and the brightness level was found. The main effects were also not statistically significant.

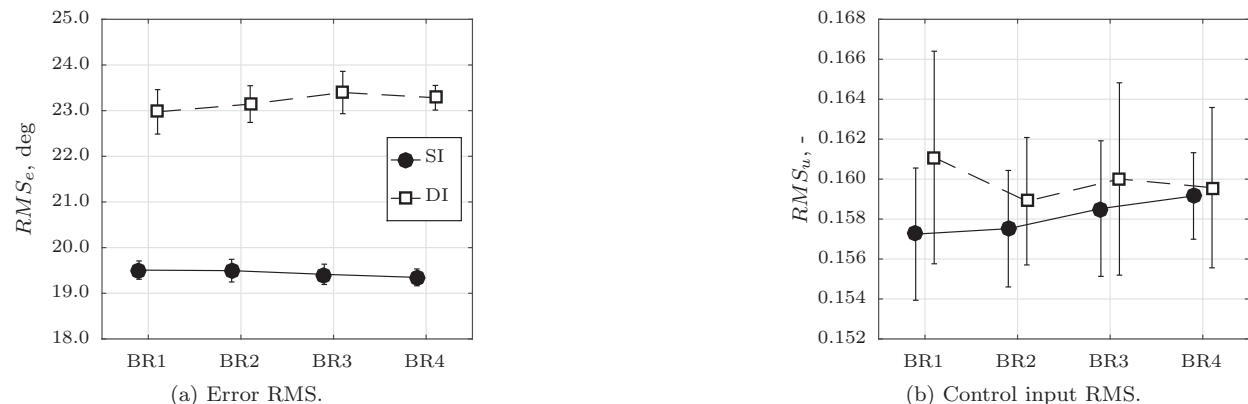
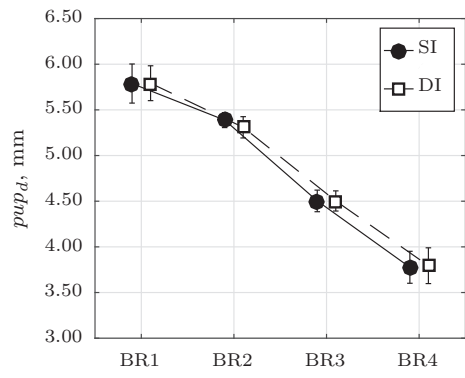


Figure 8. RMS of error and control input.

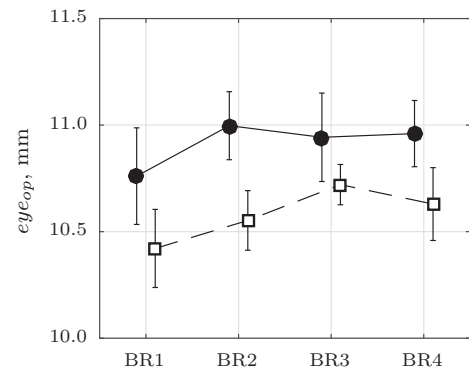
C. Eye activity parameters

Fig. 9 shows the evolution of all the eye parameters collected from the eye tracker. Pupil diameter is depicted in Fig. 9a. As expected, the pupil diameter decreases with increased display brightness almost linearly. A statistically significant main effect is obtained for the brightness factor, $F(1.2, 15.6) = 122.27, p < 0.05$. There is no main effect of the controlled dynamics, nor a significant two-way interaction between the controlled dynamics and the brightness, as can be verified from Fig. 9a. Eye opening is shown in Fig. 9b. For both the SI and DI tasks, it seems that the eye opening is the smallest for the lowest brightness of the display. Moreover, in case of the more difficult DI task, the eye opening seems smaller overall, suggesting that subjects squint while controlling this task, compared to the easier SI case. Squinting might be evidence of subjects trying to focus more during the DI task. However, both trends are not statistically significant, see Table 3. Blink count and blink duration are plotted in Fig. 9c and Fig. 9d. Blink count data suggest that participants blinked more frequently in the DI task. Furthermore, blink count seems to decrease with increased display brightness for both the SI and DI tasks. This main effect of brightness is marginally significant, $F(3, 39) = 2.48, p < 0.1$. However, post-hoc tests in which the Bonferroni correction for multiple comparisons was applied, did not reveal any significant effects for different brightnesses. Fig. 9d depicts the blink duration. The trend seems to be very similar between the SI and DI tasks, the lowest level of brightness resulting in a longer blink duration and the highest brightness in the shorter blink duration. The ANOVA reveals, however, that these effects are not significant.

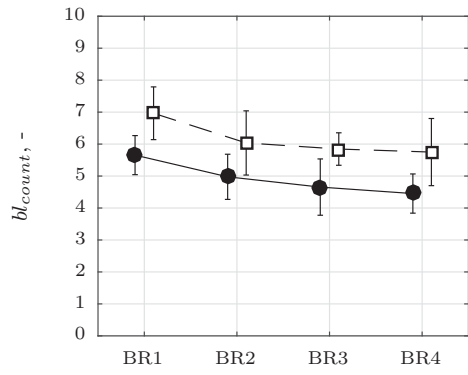
In Fig. 9, eye opening and closing amplitudes and speeds are also shown. A few interesting observations can be made. First, the eye closing speeds are higher than the opening speeds, as seen when comparing Fig. 9e to Fig. 9f. Second, opening and closing amplitudes seem to be the slowest for the lowest brightness. Lastly, the opening and closing amplitudes for the DI task are smaller compared to the SI task. This may be related to the squinting effect observed in the eye opening parameter in Fig. 9b. When the distance between the eyelids is smaller (smaller eye opening), the eyelids have to travel a shorter distance (smaller opening and closing amplitudes). However, the ANOVA analysis does not show any significant effects of controlled dynamics, brightness or the interaction between both factors for any of these dependent measures, see Table 3.



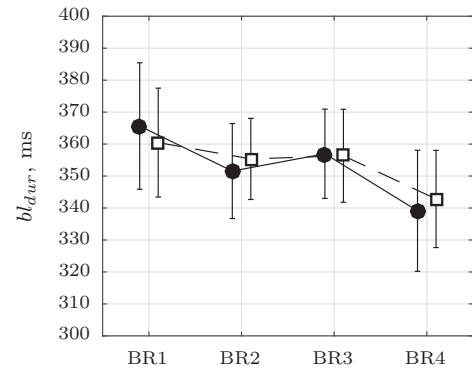
(a) Pupil diameter.



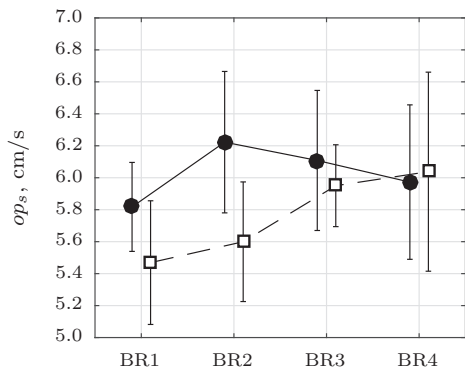
(b) Eye opening.



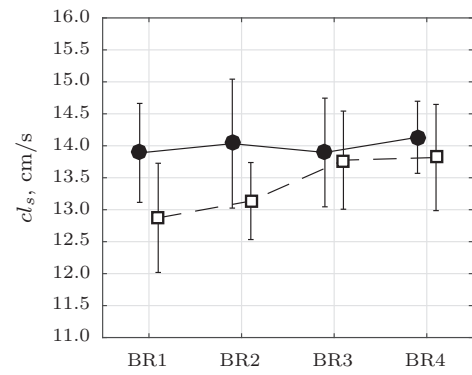
(c) Blink count.



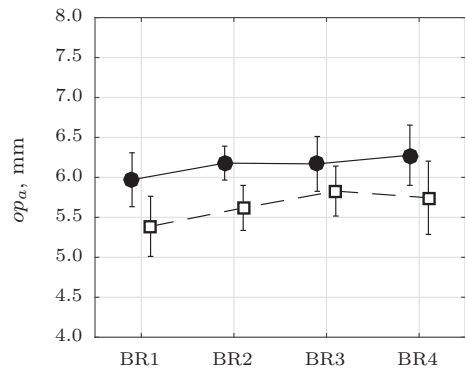
(d) Blink duration.



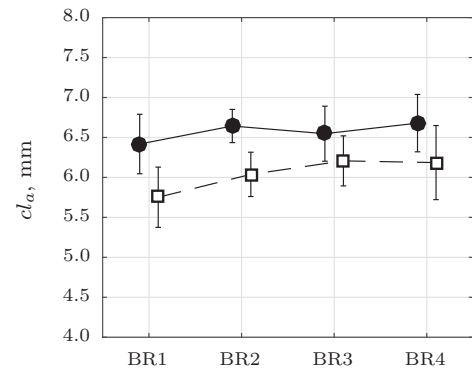
(e) Opening speed.



(f) Closing speed.



(g) Opening amplitude.



(h) Closing amplitude.

Figure 9. Eye activity measures.

D. Remnant characteristics.

Remnant power at the perceptual level is shown in Fig. 10a. No significant two-way interaction was found for this dependent measure, nor for the brightness main effect. The main effect of the controlled dynamics was statistically significant, $F(1.0, 12.0) = 8.26, p < 0.05$, the normalized remnant power being higher for the DI dynamics. Remnant gain also shows a significant difference for the main effect of dynamics, $F(1, 12) = 34.06, p < 0.05$, and no significant two-way interaction or main effect of brightness. The average remnant gain was higher for the DI case. As expected, since the remnant is highly dependent on the controlled dynamics, the remnant filter break frequency shows a significant main effect of the controlled dynamics, with $F(1, 12) = 66.63, p < 0.05$, where the break frequency is lower for the DI dynamics. The variability of the remnant break frequency is higher for the SI task. Moreover, the second brightness seems to introduce a slightly higher variability in the remnant break frequency in the SI task, compared to the other levels of brightness.

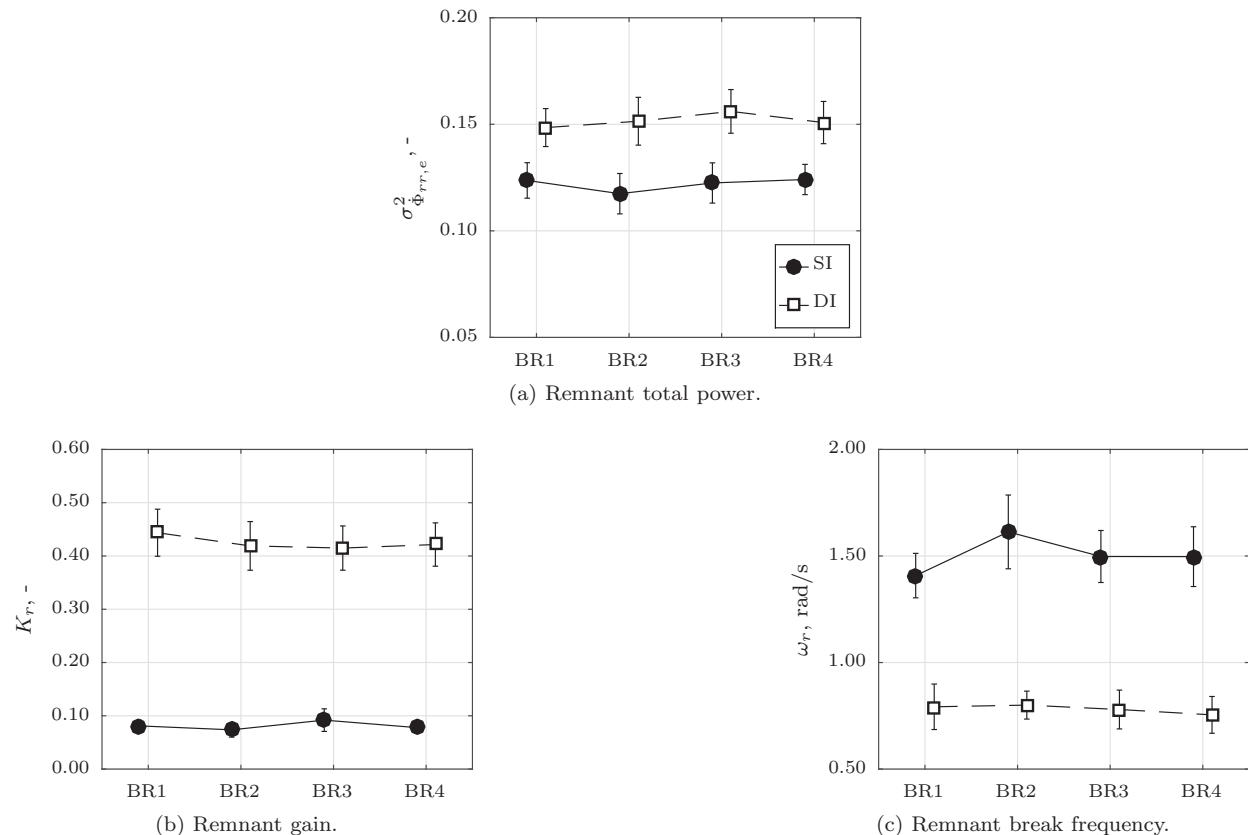


Figure 10. Remnant parameters.

E. Human operator parameters

Human operator parameters were obtained following the method presented in Section III. The visual gain is depicted in Fig. 11a. This variable seems to be invariant to different levels of brightness. No significant two-way interaction between brightness and controlled dynamics, or a significant main effect of brightness were found. A significant main effect was identified for the controlled dynamics, $F(1, 13) = 83.755, p < 0.05$, where visual gain was significantly higher for the single integrator-like dynamics by a value of 0.96.

No significant two-way interaction, or for main effects of brightness were found for the lead time constant. The main effects of the controlled dynamics was highly significant for the lead time constant depicted in Fig. 11b, $F(1, 13) = 149.86, p < 0.05$. The lead time constant is notably larger for the DI task. According to McRuer's crossover model,¹ this is expected since the human controller has to generate lead at lower frequencies compared to the SI task. For the single integrator-like dynamics, the time constant was around 0.166 seconds, meaning that the human operator is expected to generate lead at $\omega_d = 6$ rad/s. Results show that participants generated lead at around $1/0.22 = 4.54$ rad/s, at slightly lower frequencies than expected,

indicating the generation of more phase at mid-frequencies. In the case of the DI task, ω_d was 1 rad/s, and the participants generated lead at around $1/0.9 = 1.11$ rad/s, at a slightly higher frequency. In Fig. 11b, the solid horizontal line represents the break frequency of the SI dynamics, and the dashed horizontal line the break frequency of the DI dynamics.

Measured effective time delay in the control loop is depicted in Fig. 11c. Two interesting observations can be made. First, the time delay seems to be affected by display brightness for the SI task, and not for the DI task. Moreover, it is lower for the SI task compared to the DI task for all brightnesses except BR1, the darkest display background. For this variable, there was a significant two-way interaction between controlled dynamics and display brightness, $F(3, 39) = 3.99, p < 0.05$. Therefore, simple main tests were performed. First, simple main effects of dynamics were tested, at each level of display brightness. Visual time delay was not significantly different between the SI and DI tasks for brightness 1. However, it was around 16 ms lower for the SI task compared to the DI task for brightness levels 2 and 3, a statistically significant difference, $p < 0.05$. A marginally statistically significant difference ($F(1, 13) = 3.72, p < 0.1$) existed at BR4 between the time delay in the SI task compared to the DI task, an absolute difference of 11 ms. Next, simple main effects of brightness were evaluated. As expected from evaluating Fig. 11c, there was a statistically significant difference in the time delay in the SI task over the levels of brightness, $F(3, 39) = 5.42, p < 0.05$, and no statistically significant difference in the DI task. More specifically, the time delay in the SI task for BR2 was 10 ms lower compared to BR1, $p < 0.05$. Also, the time delay in the SI task for brightness level 3 was 9 ms lower compared to brightness level 1, $p < 0.1$. No other significant differences were found.

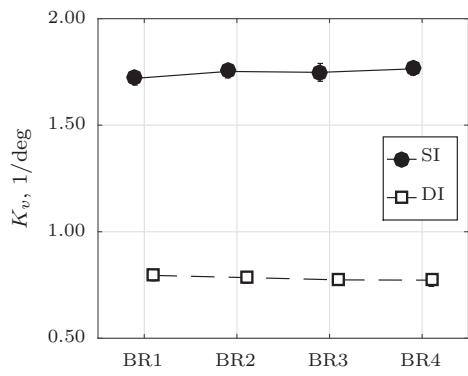
The neuromuscular damping ratio is depicted in Fig. 11d. No statistically significant two-way interaction or main effect of brightness were found for this variable. However, the main effect of the controlled dynamics showed that there was a statistically significant difference between the SI and DI conditions, $F(1, 13) = 16.51, p < 0.05$. The double integrator-like dynamics resulted in a larger neuromuscular damping compared to the DI task. Neuromuscular frequency is shown in Fig. 11e. A marginally statistically significant two-way interaction was found, $F(3, 39) = 2.316, p = 0.09$. Simple main effects of dynamics showed statistically significant differences at all levels of brightness, where the neuromuscular dynamics was higher for the SI task. Next, simple main effects of brightness were investigated. A marginally statistically significant difference in neuromuscular frequency was found for the SI task over the different levels of brightness, $F(1.8, 23) = 0.21, p < 0.1$, given by the "bowl"-like shape that can be observed in Fig. 11e. No significant difference was found for the DI task. With the single integrator-like dynamics, the neuromuscular frequency for BR2 was 0.18 rad/s smaller than for BR1, $p < 0.05$. No statistically significant differences were found for the other levels of brightness.

F. Variance Accounted For

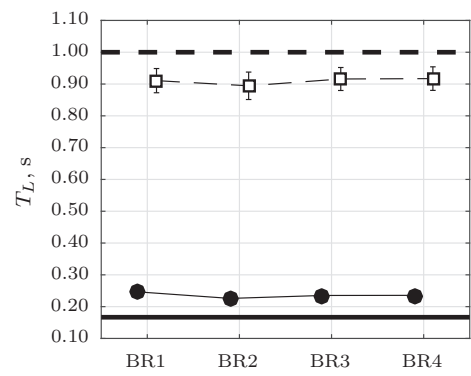
The variance accounted for indicates how much of the control input signal measured from the participants can be explained by the proposed linear model described in Section II. It can be interpreted as the accuracy with which the linear human controller model fits the measured control signal data. The VAF values for different conditions are shown in Fig. 11f. Since six runs are averaged in the time domain to reduce the effect of the remnant in order to obtain accurate pilot model parameters, the VAF values are high for all conditions and have values around 91%. No significant interactions were found between the factors in the ANOVA.

G. Crossover frequencies and phase margins

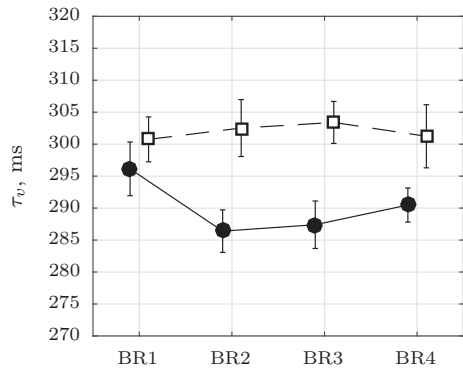
Performance and stability of the human-vehicle system are assessed by looking at the crossover frequencies and phase margins of the open-loop system, shown in Fig. 11g and Fig. 11h. The crossover frequency seems to be invariant with the controlled dynamics. The phase margins do differ however, the double integrator-like dynamics showing a reduced phase margin compared to the single integrator-like dynamics. This is probably mostly due to the inherently higher phase lag introduced in the DI task, as shown in Fig. 2b. A marginally significant two-way interaction was found for the crossover frequency, $F(2.04, 25.54) = 2.95, p < 0.1$. However, no statistically significant differences were found when running simple main effects of controlled dynamics and brightness. No significant two-way interaction between controlled dynamics and brightness was found for the phase margin. However, as suggested in Fig. 11h, there was a statistically significant difference in controlled dynamics, $F(1, 13) = 26.9, p < 0.05$. On average, the single integrator task had a 12.6 degrees higher phase margin than the double integrator task.



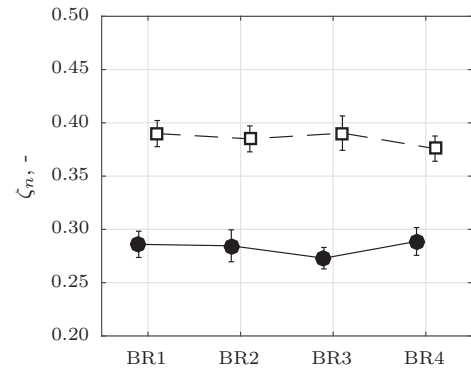
(a) Visual gain.



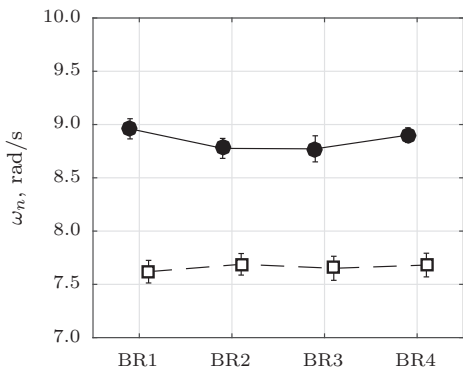
(b) Lead time constant.



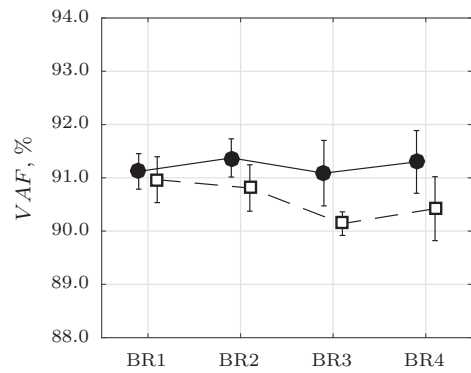
(c) Visual delay.



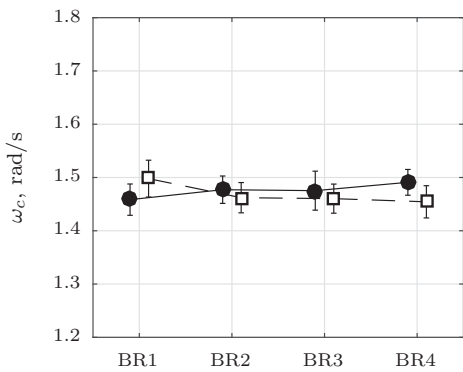
(d) Neuromuscular damping.



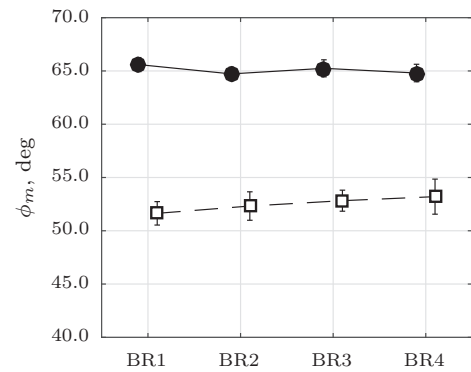
(e) Neuromuscular frequency.



(f) Variance Accounted For.



(g) Crossover frequency.



(h) Phase margin.

Figure 11. Human operator parameters, VAF, crossover frequencies and phase margins.

V. Discussion

The objective of this research was to investigate if changes in eye activity have an effect on human remnant and control behavior. Fourteen general aviation pilots performed a pitch tracking task, in which pupil diameter was controlled via different background colors of the display that impacted the perceived brightness. Both single integrator-like and double integrator-like dynamics were tested in order to verify if the possible effects depend on task difficulty.

Subjective ratings revealed that the tasks with the darkest and brightest background respectively, resulted in highest discomfort. The reason is that in condition BR1 the PFD was considerably brighter in comparison to the background, whereas in condition BR4 the background color (white) was very bright compared to the PFD in the middle.

Tracking performance, as indicated by the RMS of the tracking error, was not affected by the different levels of brightness. It was hypothesized that the tracking performance would degrade depending on how comfortable the display background was perceived by the participants (H5). However, this effect is not present in the results, as the performance was independent of brightness. With the more unstable double integrator-like dynamics, participants did show consistently larger tracking errors, thus lower performance. Control inputs were not effected by neither the display brightness nor the controlled dynamics.

Pupil diameter was the eye parameter that was purposely varied via the pupillary light reflex. As hypothesized, pupil diameter reduced almost linearly with increasing brightness (H1). Since manual tracking tasks rely on skill-based behavior of the human operator and there is no need for cognitive load, it was also hypothesized that the pupil diameter change will be identical with both dynamics (H2). It was shown that this is indeed the case, and no significant difference was introduced by the different controlled dynamics. Several other eye activity measures were recorded during the experiment. Eyelid opening seems consistently smaller in the DI task compared to the easier SI task, although the ANOVA revealed that this effect was not statistically significant. This effect might suggest, however, that participants are squinting with the more difficult task, since squinting typically helps people focus, obtaining a sharper focus by changing the shape of the cornea. Neither brightness nor controlled dynamics introduced significant variations in number of blinks or blink duration, as revealed by the ANOVA. The results suggest, however, that the more difficult task resulted in a slightly larger number of blinks, consistent for all levels of brightness. Previous studies suggest that blinks are not always random and they happen as the brain "pauses" and processes the information, disconnecting itself from the continuous intake of visual stimuli.¹⁶ It might be that with more lead required to perform the task, the brain requires more effort to process the information, which is subconsciously materialized in more blinks. Furthermore, the results also suggest that the blink duration is longer for the darkest background and shorter for the very white background, compared to the two in-between brightness levels. This might be due to the fact that the dark environment is understimulating the participants, which can result in boredom and sleepiness, reflected in longer blinks. On the other hand, it can also happen that the white background color in condition BR4 makes the blink duration shorter due to the overstimulation of the participants. Eye opening and closing amplitudes seemed smaller in the DI task for all levels of brightness, compared to the SI task. Although these results were not statistically significant, they hint again at the squinting effect encountered in the eyelid opening measure. With the eyelids closer together, it is then natural for the opening and closing movements to have smaller amplitudes. Opening and closing speeds did not significantly change with brightness or controlled dynamics, although it seems that they are slightly lower in the SI task. This difference is, however, not consistent for all levels of brightness.

Remnant characteristics were assessed in terms of total power, remnant gain, and remnant break frequency, following the model proposed by Levison et al.² The controlled dynamics used in this experiment were chosen to allow the identification of both the remnant gain and break frequency. It was hypothesized that some of the eye measures will have an effect on remnant characteristics (H4). However, none of the characteristics of the human remnant obtained at the perceptual level showed changes with the display brightness. This already shows that either perceptual remnant is invariant with the amount of light entering the eye (since pupil diameter significantly changed over different brightnesses) or, the influence of the eye activity on its characteristics is very small and could not be detected. Total remnant power was slightly higher for the DI task, a significant result. However, it did not change according to the subjective ratings, as hypothesized (H5). With the more difficult dynamics, the remnant gain was higher and the break frequency lower ($K_r = 0.42$, $\omega_r = 0.75$ rad/s) than the SI task ($K_r = 0.1$, $\omega_r = 1.5$ rad/s), as hypothesized (H3). The results for the remnant break frequency are in agreement with the findings of Jex et al.⁷ and McRuer,¹⁷ considering the controlled dynamics used in the current experiment.

Human operator visual gain was not affected by different brightness levels. Visual gain was lower in the DI task compared to the SI task, this difference being influenced by the different gains used in transfer function of the controlled dynamics. Due to the nature of the controlled dynamics, human operators had to generate lead for both dynamics. Naturally, the lead time constant was higher for the dynamics requiring more lead generation, DI. It is interesting to note that lead is generated at a lower frequency than the controlled dynamics break frequency for the DI task, and at a higher frequency for the SI task. This result is very similar to the findings of Zollner et al.,¹⁸ for whom the lead generation also happened at a later frequency. In their experiment, the difference between the lead time constant and the controlled dynamics time constant decreased with decreasing values of the dynamics time constant.

An interesting result is obtained for the human operator time delay, for which a two-way interaction between the dynamics and brightness was found. First, the time delay for the DI task is in general higher than the delay in the SI task. This is expected, since it is known that the time delay is dependent on the lead that has to be generated by the human operator.¹⁷ However, this expected effect was not found for BR1, the task with the darkest display. In this condition, the time delay is not significantly different for the two controlled dynamics. Furthermore, the time delay for BR1 in the SI task is around 10 ms higher compared to the other brightnesses. One possible explanation for this is that, since the pupil is fully opened at this brightness level, it is harder to focus on the PFD due to the large aperture that results in a larger circle of confusion on the retina. This effect is not present in the DI task probably because most of the time delay is due to the lead generation and the impact of pupil opening on the time delay in this task is negligible. Another possible explanation is that the squinting effect observed in the DI task helps the participants focus better in this task with a black display background, eliminating the effect seen in the SI task. It can also just be that the time delay is related to the subjective ratings for the SI task, where the dark and bright backgrounds resulted in the worst subjective ratings, and BR1 affected the time delay the most in reality. Again, the difficulty of DI might make the effects of brightness on the time delay negligible.

Performance and stability of the closed loop were assessed by the crossover frequency and phase margin of the human operator-controlled dynamics open-loop frequency response function. Crossover frequency was not statistically significantly affected by brightness or controlled dynamics. It is interesting to notice that the crossover frequency is virtually identical for both controlled dynamics. This is not in agreement with the findings of McRuer et al.,¹ who found a significantly higher crossover frequency for dynamics similar to the one in the DI task. There are various possible reasons for this finding. First of all, all participants in the current study were general aviation pilots with no previous tracking experience. Pilots in general tend to avoid high frequency inputs, which results in a low crossover frequency in the experiment, regardless of the control dynamics. Moreover, the bandwidth of the forcing function was small (1.21 rad/s), which was shown to affect the behavior of the crossover frequency.¹⁷ In addition, crossover frequencies that are similar for both single integrator and double integrator dynamics were found in another experiment.¹⁹ Lastly, it was previously shown that in a roll compensatory tracking task, display gain has a big impact on the crossover frequency and phase margin of the open-loop system.²⁰ More specifically, small display gains resulted in significantly lower crossover frequencies and higher phase margins. Although a pitch tracking task, it is possible that a higher display gain would have resulted in a larger crossover frequency. In the current experiment, the phase margin is significantly higher by around 13° for the SI task compared to the DI task. This is probably due to a combination of smaller phase lag and the differences between the generated lead and the break frequency of the controlled dynamics.

VI. Conclusions

Human remnant behavior is relatively well understood in terms of its behavior relative to task variables such as input forcing function and controlled dynamics. However, its possible link to physiological measures was never investigated. The main goal of the present work was to investigate the possible role of eye activity in the characteristics of human perceptual remnant. The eye parameter that was controlled was the pupil diameter, which was varied via the involuntary pupillary light reflex by changing the background color of a simplified PFD. In a pitch tracking task, two different controlled dynamics were tested, similar to a single integrator and a double integrator, in order to investigate if the possible relation between eye parameters and human remnant is affected by the task difficulty.

The only eye parameter that statistically significantly changed with different brightness levels was the pupil diameter through the pupillary light reflex, as intended. Moreover, the results suggested a squinting

effect, when subjects were controlling the more difficult dynamics, visible in eye opening and eye closing and opening amplitudes, probably in an attempt to obtain a better focus on the PFD. Remnant characteristics were invariant to the four different brightness levels and only changed with the different controlled dynamics.

A few interesting findings are worth mentioning. First, although subjects reported high levels of discomfort caused by the very dark and very bright background (revealed in the subjective ratings), tracking performance, together with closed-loop performance and stability were unaffected by the different levels of brightness. Second, a higher time delay for the single integrator-like task with the very dark background was found, an effect not visible with the more difficult dynamics. This might suggest that human operator limitations, and not performance, are affected by environment variables (brightness). It is possible that, with the task becoming more difficult, the environmental effects become negligible and task difficulty becomes the only responsible for such limitations. Future experiments should investigate the physiological ways in which the humans adapt to more difficult tasks, reflected in the human operator model parameters.

No definite conclusions can be drawn regarding the relation between eye activity and human perceptual remnant. The only eye parameter controlled in the experiment was the pupil diameter, and it was shown that the remnant is not related to the amount of light entering the eye, but it might be affected by other eye measures. However, since no big changes were observed in other eye parameters, it is not yet possible to draw any conclusions. Another possibility is that eye activity has a very small effect on the remnant characteristics, and the amount of data collected lacks the statistical power to reveal any significant interactions.

It could be that, rather than eye activity parameters affecting the remnant characteristics and human operator control behavior, the controlled dynamics affects the eye parameters, in turn. Although not statistically significant, the more difficult task resulted in a possible squinting effect and changed the remnant power and parameters. More subjects will be tested in the future in order to verify the existence of this phenomenon. If this is the case, then monitoring the eye activity could represent a method of predicting changes in human remnant characteristics caused by the level of task difficulty.

References

- ¹McRuer, D. T. and Jex, H. R., "A review of quasi-linear pilot models," *Human Factors in Electronics, IEEE Transactions on*, Vol. 3, 1967, pp. 231–249.
- ²Levison, W. H., Baron, S., and Kleinman, D. L., "A Model for Human Controller Remnant," *Man-Machine Systems, IEEE Transactions on*, Vol. 10, No. 4, Dec 1969, pp. 101–108.
- ³Flach, J. M., "Control With an Eye for Perception: Precursors to an Active Psychophysics," *Ecological Psychology*, Vol. 2, No. 2, 1990, pp. 83–111.
- ⁴Pew, R., Duffendack, J., and Fensch, L., "Sine-Wave Tracking Revisited," *Human Factors in Electronics, IEEE Transactions on*, Vol. HFE-8, No. 2, June 1967, pp. 130–134.
- ⁵McRuer, D. T., Graham, D., Krendel, E., and Reisener Jr, W., "Human pilot dynamics in compensatory systems: theory, models and experiments with controlled element and forcing function variations," *Wright-Patterson AFB (OH): Air Force Flight Dynamics Laboratory*, 1965.
- ⁶Levison, W. H., Elkind, J. I., and Ward, J. L., *Studies of multivariable manual control systems: A model for task interference*, National Aeronautics and Space Administration, 1971.
- ⁷Jex, H. R. and Magdalen, R. E., "Corroborative Data on Normalization of Human Operator Remnant," *Man-Machine Systems, IEEE Transactions on*, Vol. 10, No. 4, Dec 1969, pp. 137–140.
- ⁸Ellis, C., "The pupillary light reflex in normal subjects." *British Journal of Ophthalmology*, Vol. 65, No. 11, 1981, pp. 754–759.
- ⁹Watson, A. B. and Yellott, J. I., "A unified formula for light-adapted pupil size," *Journal of Vision*, Vol. 12, No. 10, 2012, pp. 12–12.
- ¹⁰Pool, D. M., Zaal, P. M. T., Damveld, H. J., Van Paassen, M. M., and Mulder, M., "Evaluating Simulator Motion Fidelity using In-Flight and Simulator Measurements of Roll Tracking Behavior," *Proceedings of the AIAA Modeling and Simulation Technologies Conference*, 2012, pp. 1–25.
- ¹¹Damveld, H. J., Beerens, G. C., Van Paassen, M. M., and Mulder, M., "Design of forcing functions for the identification of human control behavior," *Journal of Guidance, Control, and Dynamics*, Vol. 33, No. 4, 2010, pp. 1064–1081.
- ¹²Hossain, G. and Yeasin, M., "Understanding effects of cognitive load from pupillary responses using hilbert analytic phase," *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, 2014, pp. 375–380.
- ¹³Zaal, P. M. T., Pool, D. M., Chu, Q. P., Mulder, M., Van Paassen, M. M., and Mulder, J. A., "Modeling human multimodal perception and control using genetic maximum likelihood estimation," *Journal of Guidance, Control, and Dynamics*, Vol. 32, No. 4, 2009, pp. 1089–1099.
- ¹⁴Van Paassen, M. M. and Mulder, M., "Identification of human operator control behaviour in multiple-loop tracking tasks," *Proceedings of the Seventh IFAC/IFIP/IFORS/IEA Symposium on Analysis, Design and Evaluation of Man-Machine Systems*, 1998, pp. 515–520.
- ¹⁵Mulder, M., Mulder, J. A., and Stassen, H. G., "Cybernetics of tunnel-in-the-sky displays. I. Straight trajectories,"

Systems, Man, and Cybernetics, 1999. IEEE SMC99 Conference Proceedings. 1999 IEEE International Conference on, Vol. 5, IEEE, 1999, pp. 1082–1087.

¹⁶Nakano, T., Kato, M., Morito, Y., Itoi, S., and Kitazawa, S., “Blink-related momentary activation of the default mode network while viewing videos,” *Proceedings of the National Academy of Sciences*, Vol. 110, No. 2, 2013, pp. 702–706.

¹⁷McRuer, D. T., “Human dynamics in man-machine systems,” *Automatica*, Vol. 16, No. 3, 1980, pp. 237–253.

¹⁸Zollner, H. G. H., Pool, D. M., Damveld, H. J., Van Paassen, M. M., and Mulder, M., “The effects of controlled element break frequency on pilot dynamics during compensatory target-following,” *AIAA Guidance, Navigation, and Control Conference and Exhibit*, 2010.

¹⁹Lu, T., Pool, D., Van Paassen, M. M., and Mulder, M., “Use of simulator motion feedback for different classes of vehicle dynamics in manual control tasks,” *Proceedings of the 5th CEAS Air and Space Conference; Challenges in European Aerospace, Delft (The Netherlands) 7-11 Sept. 2015*, Delft University of Technology, 2015.

²⁰Breur, S. W., Pool, D. M., Van Paassen, M. M., and Mulder, M., “Effects of Displayed Error Scaling in Compensatory Roll-Axis Tracking Tasks,” *AIAA Modeling and Simulation Technologies Conference, Toronto, Canada, 2-5 August 2010; AIAA 2010-8091*, American Institute of Aeronautics and Astronautics (AIAA), 2010.

²¹Klingner, J., Kumar, R., and Hanrahan, P., “Measuring the task-evoked pupillary response with a remote eye tracker,” *Proceedings of the 2008 symposium on Eye tracking research & applications*, ACM, 2008, pp. 69–72.

A. Description of eye parameters recorded during the experiment

Eye opening parameters

In Fig. 12, the time evolution of the eyelid opening recorded by the Smart Eye Pro system before, during blink is shown. This gives an insight on how some of the eye parameters are obtained during the experiment by the eye tracker.

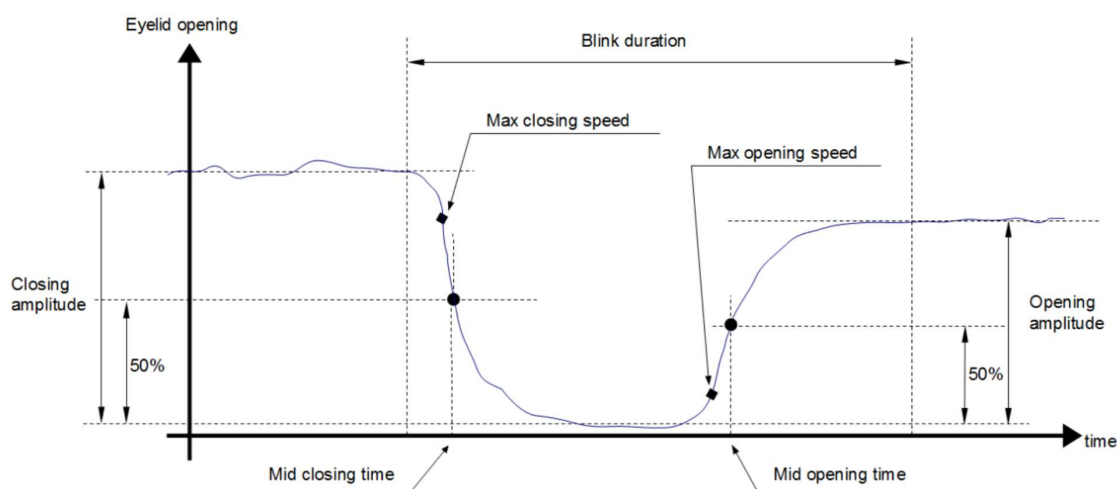


Figure 12. Explanation of some of the eye parameters. Adapted from the Smart Eye Pro manual.

Eyelid opening is defined by the distance between the top eyelid and the bottom eyelid of the participant, as measured by the eye tracker system. The eye closing amplitude is the difference between the eyelid opening before the blink and the minimum point in the eyelid opening value. The eye opening amplitude is determined similarly. Note that the two measures are usually not equal. The opening and closing speeds represent the time derivative of the eyelid opening value, after the eye starts closing until it is almost closed, and from when it starts opening until it is fully opened again, respectively. Blink duration is calculated as the duration between the eyelid starting dropping until it reaches a stable non-zero level again. Movements similar to blinks which last longer than 700 ms are not considered as blinks by the eye tracker system.

Pupil diameter filtering

The pupil diameter was one of the continuous variables recorded during the experiment. A third order Butterworth filter was applied in order to eliminate system noise. In order to choose the break frequency of this filter, a method based on correlation of the data from the left and right eyes was used, as follows:

- A third order Butterworth band-pass filter with a bandwidth of 1 Hz was applied to the pupil diameter measurements from both the left and right data; the low and high frequencies were initially set to 0.01 and 1.01 Hz respectively, the zero frequency being located at 0.5 Hz.
- The correlation coefficient between the two resulting signals is calculated
- The band-pass filter zero frequency is moved in steps of 0.5 Hz, and the correlation coefficient is calculated again.
- The calculation stops when the zero frequency of the band-pass filter reaches 25 Hz.

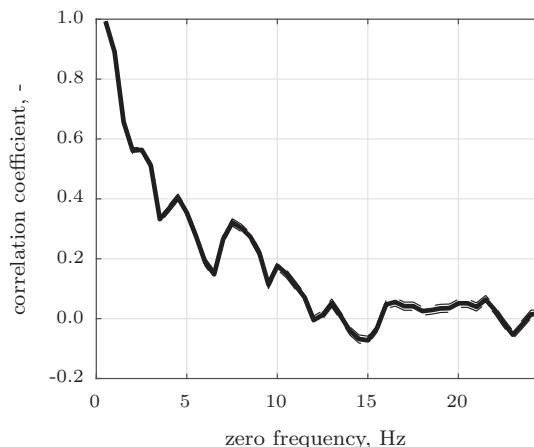


Figure 13. Correlation coefficient between the left and right eye, at different frequencies.

A similar method was successfully applied by Klingner et al.,²¹ in a research aimed at analyzing task-evoked pupillary responses. Fig. 13 shows an example of the correlation coefficient method obtained from one subject. Correlation coefficient is placed on the vertical axis and the zero frequency of the band-pass filter on the horizontal axis. As can be seen in Fig. 13, at around 12 Hz the correlation coefficient becomes negative, meaning that the combined information from the left and right eyes is not usable above this frequency. Multiple tests with different subjects were performed, and the average frequency at which the correlation coefficient becomes negative was between 10 and 15 Hz. Therefore, to be on the conservative side, a third order Butterworth low-pass filter with a break frequency of 10 Hz was chosen for filtering the pupil diameter in the experiment.

B. Description of the applications used in the experiment

This appendix provides an overview of the applications used in the experiment, where their interaction and properties are further explained.

Fig. 14 is a schematic view of the two different systems used in the experiment. The first one, running at 100 Hz, is used for the pitch tracking task, and the other is connected to the Smart Eye Pro system collecting data at 60 Hz. Ideally, the time used for recording data on both machines is the same. The two PC systems were connected by cable via a router. Since the tracking task application was developed in Python, a Smart Eye API for Python was used in order to send commands to the eye tracker application.

The difference in sampling frequency from the two systems did not pose many problems for the purpose of the experiment, since the mean eye tracker data from one run of 81.92 seconds were used in the experiment. The only important aspect was for the data to have the same start and end time used for recording. When the pitch tracking task started, a command was sent via the UDP connection from the tracking task application to the Smart Eye system to start tracking and logging the data. Both machines log the Unix epoch time in milliseconds. After running multiple tests, it was found that the delay between the two systems was invariant to the length of the run and it varied between 300 - 1000 milliseconds. This is probably due to the fact that it takes some time for the eye tracker system to start tracking and log the data. In order to deal with this delay, an extra three seconds were added to the beginning of the runtime, and the time domain data were aligned using the logged Unix epoch time.

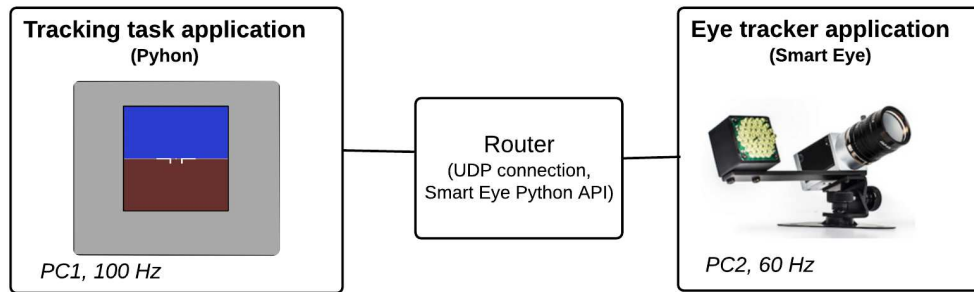


Figure 14. Schematic view of the systems used in the experiment.

Fig. 15 shows the user interface of the tracking task, including the PFD that the participants saw during the experiment.

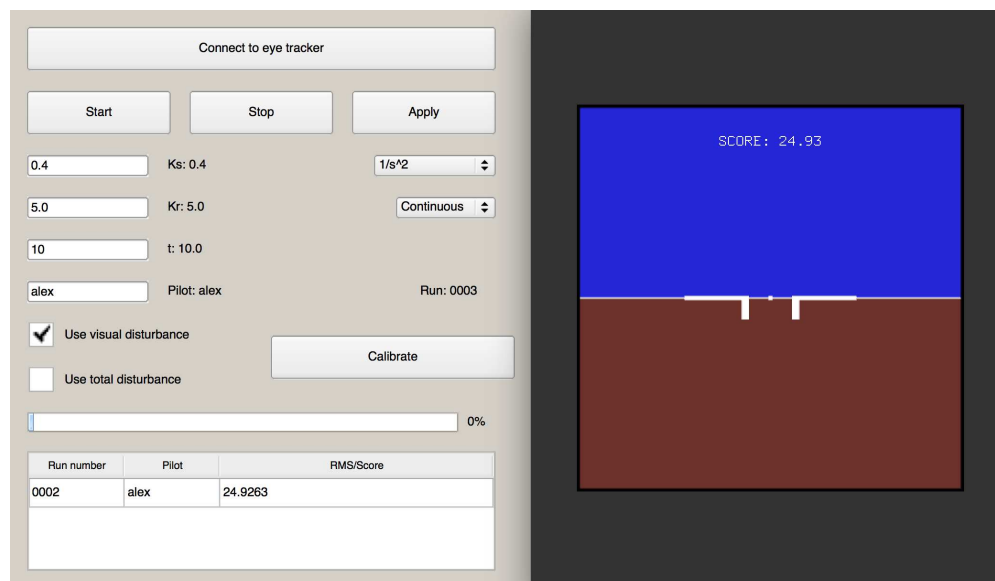


Figure 15. User interface of the tracking task application.

The main goal of the user interface was to set parameters used in the tracking task experiment. The parameters that can be set are: stick gain ("Ks"), display gain ("Kr"), participant name ("Pilot"), runtime ("t"), controlled dynamics (single integrator, double integrator, somewhere in-between), forcing functions (target, disturbance, or both). Furthermore, "continuous" indicates that there will only one run performed. The other option, "with pauses", will run the task eight times, with 15 seconds break between each run. Moreover, the bottom of the user interface shows the run number for the participant with the given name and his performance, indicated by the RMS of the tracking error. The button "Connect to eye tracker" attempts to establish the connection with the eye tracker system. On the right side of Fig. 15, the PFD is shown, with the RMS score at the end of the run displayed. When the "Start" button is pressed, the PFD is maximized and the tracking task starts after three seconds, as mentioned before.

In order to give a more detailed overview of the logic behind the tracking task application, Fig. 16 shows its different components and their roles in the application.

In the *main.py* file, the user interface shown in Fig. 15 is defined. Apart from containing the user interface and initializing all the parameters, including the joystick, gains, runtime, controlled dynamics, connection with the eye tracker, it also represents the center that communicates with all the other modules. After the

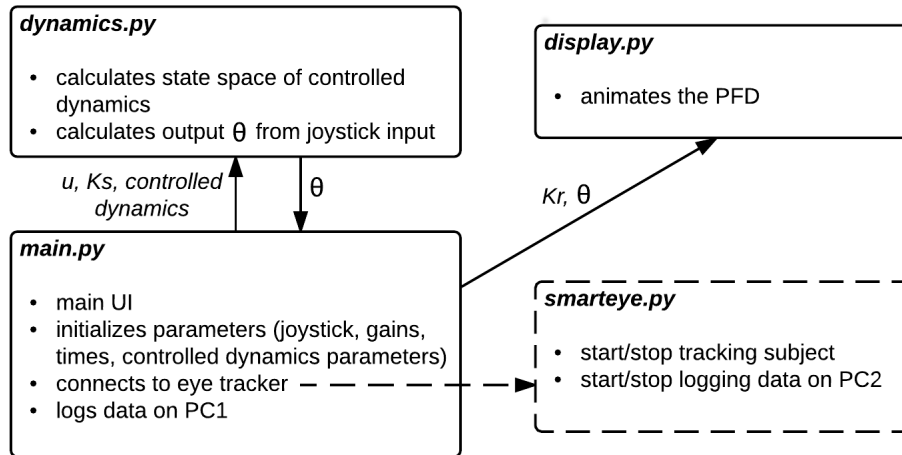


Figure 16. Components of the tracking task application and their interaction.

user selects the stick gain and desired controlled dynamics, the state space representation of the controlled dynamics is created in *dynamics.py*. The joystick input value u is then used to calculate the output pitch angle θ , which is returned to the main file. The main file then commands the *display.py* file to animate the PFD, using the calculated θ and display gain "Kr". Furthermore, the main file makes the connection between the two PCs through the Smart Eye API for python. This file sends the start/stop commands for eye tracking and logging commands. Finally, the main file logs all the measurable signals in the control loop on the tracking application machine.

C. ANOVA results

This appendix provides a summary of the two-way repeated measures ANOVA on all the dependent measures. P values below 0.05 are considered highly significant (**), between 0.05 and 0.1 marginally significant (*), and above 0.1 not significant (-).

Table 3. Summary of the two-way repeated measures ANOVA.

Dependent Variable	DYNxBR				DYN				BR			
	df	F	p	Sig.	df	F	p	Sig.	df	F	p	Sig.
RMS_e	3, 39	1.187	0.327	–	1, 13	216.035	0.000	**	2.117, 27.516	0.247	0.795	–
RMS_u	3, 39	0.416	0.742	–	1, 13	0.043	0.839	–	3, 39	0.103	0.958	–
N_{POW}	3, 36	0.188	0.904	–	1, 12	8.262	0.014	**	3, 36	0.431	0.732	–
K_r	3, 36	0.292	0.831	–	1, 12	34.057	0.000	**	3, 36	0.342	0.795	–
ω_r	3, 36	0.987	0.410	–	1, 12	66.632	0.000	**	3, 36	0.528	0.666	–
pup_d	3, 39	0.856	0.472	–	1, 13	0.125	0.729	–	1.196, 15.554	122.267	0.000	**
bl_{count}	3, 39	0.053	0.984	–	1, 13	1.908	0.190	–	3, 39	2.477	0.076	*
bl_{dur}	3, 36	0.165	0.919	–	1, 12	0.018	0.895	–	3, 36	1.923	0.143	–
eye_{op}	3, 39	1.024	0.393	–	1, 12	0.845	0.375	–	1.721, 22.378	1.807	0.190	–
op_s	1.928, 23.137	0.960	0.395	–	1, 12	0.888	0.365	–	3, 36	1.099	0.362	–
op_a	3, 36	0.277	0.842	–	1, 12	1.959	0.187	–	3, 36	1.072	0.373	–
cl_s	3, 36	0.482	0.697	–	1, 12	0.722	0.412	–	3, 36	0.897	0.452	–
cl_a	3, 36	0.517	0.673	–	1, 12	1.963	0.186	–	3, 36	0.994	0.406	–
K_v	3, 39	1.741	0.175	–	1, 13	83.755	0.000	**	3, 39	0.218	0.884	–
T_L	3, 39	0.159	0.923	–	1, 13	149.858	0.000	**	3, 39	0.619	0.607	–
τ_v	3, 39	3.985	0.014	**	1, 13	5.277	0.039	**	3, 39	1.260	0.301	–
ζ_n	3, 39	1.871	0.150	–	1, 13	16.507	0.001	**	3, 39	0.299	0.826	–
ω_n	3, 39	2.316	0.090	*	1, 13	68.890	0.000	**	3, 39	1.015	0.397	–
VAF	3, 39	1.268	0.299	–	1, 13	1.057	0.323	–	3, 39	1.277	0.296	–
ω_c	2.042, 26.541	2.947	0.069	*	1, 13	0.008	0.930	–	3, 39	0.125	0.945	–
ϕ_m	3, 39	1.515	0.226	–	1, 13	26.900	0.000	**	3, 39	0.447	0.721	–