# System Delay in Flight Simulators Impairs Performance and Increases Physiological Workload

Nina Flad[1], Frank M. Nieuwenhuizen[1], Heinrich H. Bülthoff[1,2], and Lewis L. Chuang[1,⋆]

[1] Department of Perception, Cognition and Action,
Max Planck Institute for Biological Cybernetics, Tübingen
[2] Department of Cognitive and Brain Engineering, Korea University
`{nina.flad,frank.nieuwenhuizen,heinrich.buelthoff,`
`lewis.chuang}@tuebingen.mpg.de`

**Abstract.** Delays between user input and the system's reaction in control tasks have been shown to have a detrimental effect on performance. This is often accompanied by increases in self-reported workload. In the current work, we sought to identify physiological measures that correlate with pilot workload in a conceptual aerial vehicle that suffered from varying time delays between control input and vehicle response. For this purpose, we measured the skin conductance and heart rate variability of 8 participants during flight maneuvers in a fixed-base simulator. Participants were instructed to land a vehicle while compensating for roll disturbances under different conditions of system delay. We found that control error and the self-reported workload increased with increasing time delay. Skin conductance and input behavior also reflect corresponding changes. Our results show that physiological measures are sufficiently robust for evaluating the adverse influence of system delays in a conceptual vehicle model.

## 1 Introduction

Delays between input and feedback in a closed-loop control task can result in both perceptual and control instabilities. For example, in head-slaved visualization systems (i.e., head-mounted virtual reality displays), temporal discrepancies between head movements and display updating can result in oscillopsia (also referred to as simulator sickness) in which the human operator perceives an unstable virtual environment that "swims around" his head [1]. In vehicle simulators, time delays between manual inputs and visual feedback can lead to notable increases in performance errors as well as perceived workload [2–4]. The latter can induce stress that induces physiological reactions in the autonomic nervous system.

---

[⋆] The work in this paper was supported by the myCopter project, funded by the European Commission under the 7th Framework Program.

D. Harris (Ed.): EPCE 2014, LNAI 8532, pp. 3–11, 2014.

Previous studies of flight performance have shown that visual feedback delays can decrease performance and increase workload. For example, participants who performed a low-level flight task under visual lag conditions produced larger altitude errors and responded with higher workload scores on a questionnaire [5]. In a different study, increasing system delays decreased piloting performance as well as the subjective handling qualities of the aircraft, when pilots were required to perform a side-step maneuver in a helicopter simulation as well as actual test flight [4].

This decrease in performance can be more specifically attributed to the influence of visual feedback delays on closed-loop control performance. Trying to compensate for a time-delayed error has been shown to result in pilot-induced oscillations, wherein the control inputs from the pilot actually adds to the overall system disturbance instead of subtracting from it [2]. This is especially detrimental to the performance of precision tasks, such as hovering or landing. If the pilot is trained on such maneuvers in a simulator that suffers from time delays, more time is necessary to acquire the targeted skill and the transfer of training to real flight could be problematic, since a real aircraft with no system delays can be expected to respond differently [6]. Moreover, training with system delays has also been shown to increase the workload that is subjectively experienced by the pilot [5].

Many studies employ questionnaires to assess the participants' workload. Although this approach is well established, it has known weaknesses; the measurement is obtrusive and cannot be conducted during the task itself. Therefore, self-assessment of workload relies on the participant's recollection of the task, which could be subjectively altered.

An increase in workload can induce stress, which in turn leads to psychophysiological reactions of the autonomic nervous system. For example, induced workload can increase the heart rate as well as disrupt the regular fluctuations of inter-heartbeat-intervals [7]. In addition, it can widen the perspiratory glands and affect skin conductance [8]. Both heart activity and skin conductance can be measured using skin electrodes, thus providing an online metric for stress and workload during control activity itself.

In the current work, we investigate the influence of system delays on the control of a personal aerial vehicle (PAV) concept model [9]. We introduced delays of 0, 200, 400 or 600 ms and the influence of these delays was assessed in terms of our participants' control performance, control inputs, physiological responses and questionnaire results.

The remainder of this paper is organized as follows. Section 2 describes the flight task as well as the simulator and the aircraft model, followed by a description of data acquisition and analysis. Section 3 presents the results and possible interpretations. In section 4 we summarize our findings and discuss the implications for flight simulator studies.

**Table 1.** The parameters for the disturbance in the roll axis of the vehicle

| i | $a_i$ | $\psi_i$ | $f_i$ |
|---|---|---|---|
| 1 | 0.5 | 0 | 0.0159 |
| 2 | 0.3 | 1 | 0.0796 |
| 3 | 0.2 | 0 | 0.0477 |
| 4 | −0.2 | 1 | 0.0159 |

## 2 Methods

### 2.1 Participants

Eight male participants took part in our study. Their ages ranged between 22 and 34 years. All were researchers at the Max Planck Institute for Biological Cybernetics and had normal or corrected-to-normal vision.

### 2.2 Apparatus and Flight Model

We evaluated the effect of system delay of a dynamic PAV concept model in a fixed-base flight simulator. The simulator consisted of a display wall of nine screens taking a field-of-view of 105° by 100°. The participants used generic helicopter controls consisting of foot pedals, collective stick and cyclic stick.

The outside visualization was provided by Flightgear, an open-source flight simulator [10], while the control model was implemented in Matlab/Simulink and running at 256 Hz. The control model simulated the vehicle's dynamics and calculated the position and orientation of the aircraft based on the current control inputs. The outputs were then transformed into world coordinates and sent to the Flightgear computers that rendered the scene, namely San Francisco International Airport. The landing zone measured approximately 55 by 260 meters and was at the end of a runway (see Figure 1).

The PAV model represents an augmented helicopter with uncoupled cyclic, collective and pedal controls. Its rotational dynamics were of the Attitude Command-Attitude Hold (ACAH) type, such that a constant deflection of the cyclic stick resulted in a constant rotational attitude. Participants directly controlled a rate in the heave axis with the collective stick. A constant input on the pedals resulted in a specific rotational rate around the yaw axis. Subjects had full control over all the vehicle's degrees of freedom during each trial. In our experiment, a time delay of 0, 200, 400 or 600 msec was introduced between the control input and vehicle dynamics. These values were chosen based on a pilot study.

In addition, a disturbance was introduced in the roll axis during flight. Thus, our participants had to compensate for this disturbance even whilst performing the primary task of landing the PAV. The forcing function was a summation of four sinusoidal functions

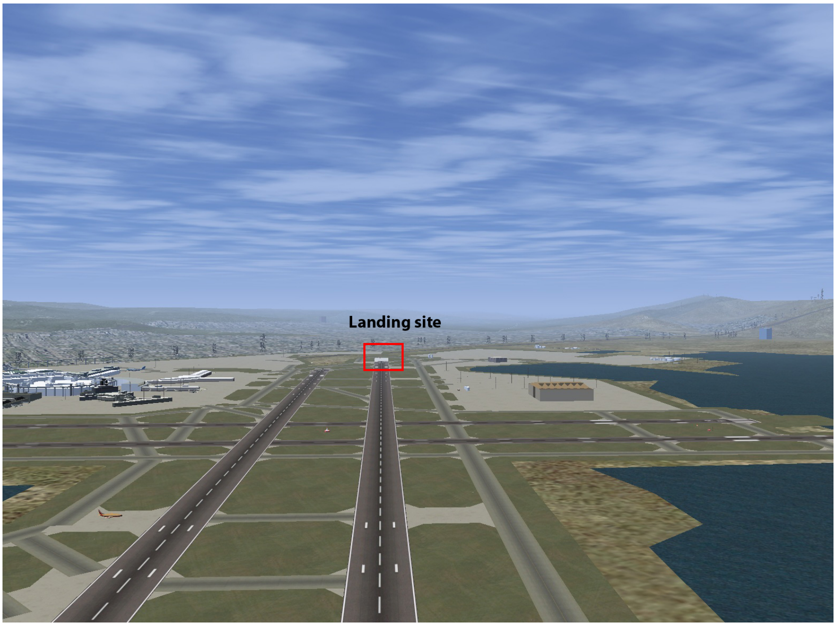

**Fig. 1.** Landing site as seen from the participants upon trial initiation

$$d(t) = \sum_{i=1}^{n} a_i * \sin(\psi_i * \frac{\pi}{2} + 2\pi * f_i * t) \tag{1}$$

with the parameters $a_i$, $\psi_i$ and $f_i$ given in Table 1.

### 2.3 Procedure

Our participants were instructed to fly the PAV from their initial position at the start of the trial towards the visible airport, follow the runway and land in a designated area at the end of the runway. In addition, they were required to maintain PAV stability and to compensate for disturbances in roll axis. Upon a successful landing, they were required to press a button to end the trial. Alternatively, each trial ended automatically if the maneuver was not successfully completed within eight minutes.

Prior to data collection, every participant had at least five one-hour training sessions with the simulator and the PAV model. During the first two training sessions, there were neither disturbance nor system delay to facilitate user familiarization with the control devices and the vehicle's dynamics. In addition, participants had to learn to navigate by relying on visual landmarks near the landing site. In the next two sessions, the roll disturbance was introduced, but without any time delay. Each training session always consisted of five flight sessions, separated by a thirty second break.

After the first four training sessions, the participants experienced at least one additional session under actual experimental conditions. The sessions for data collection consisted of four trials that varied in roll disturbance and time delay (0 ms, 200 ms, 400 ms and 600 ms) separated by a break of five minutes. In this break, participants were asked to rate workload with a digital version of the NASA-TLX questionnaire on a separate laptop. We collected ECG and skin conductance values during flight as well as during the breaks.

### 2.4 Data Collection and Analysis

Subjective workload was assessed using a computerized NASA Task Load Index (NASA-TLX). This rating scale consists of six independent scales, defined as follows [11]:

- Mental Demand (e.g. thinking, deciding, remembering, looking, etc.)
- Physical Demand (e.g. pushing, pulling, activating, etc.)
- Temporal Demand (e.g. time pressure)
- Performance
- Effort (required to achieve the level of performance)
- Frustration Level

This questionnaire was administered after every condition. It provides an overall workload score as well as scores for each individual scale and the composition of the overall score by the individual scales.

To measure performance, we calculated the root mean square error in compensating for the roll disturbances as well as control input activity. When participants experience a subjective loss of control or are performing badly, they tend to alter their input behavior or control effort. Therefore, we analyzed changes in the stick input activity. The stick input was collected at 256 Hz and analyzed in the frequency domain. We evaluated spectral densities in the frequencies higher than 0.1 Hz, since the disturbances took place in frequencies lower than 0.08 Hz. Thus, any changes in bandwidth above 0.1 Hz could be attributed to the pilot and not our disturbance function.

The first physiological measure is the skin conductance, which can be measured with a constant potential. The human skin possesses a natural electrical resistance, but contains sweat glands serving as conductive channels. Higher activity in the sweat glands results in lower resistance and better conductance [12]. The sweat glands are innervated by sympathetic activity only and, therefore, the skin conductance can serve as an indicator for stress and anxiety [8]. For the analysis, we normalized the mean conductance of each trial to a baseline measurement taken before the first test trial.

The second physiological measure is based on ECG measurements. The mean heart rate changes constant in response to changing environmental demands. These changes occur periodically and depend on the mental and physical state of the human. They are evoked by activity in the (para)sympathetic nervous system and have been found to be sensitive to work conditions, such as before and after a driving task [7], or different phases of a monitoring and detection experiment [13]. We collected ECG data at 256 Hz and filtered it offline. The heart-beats in this signal were detected and the instantaneous heart-rate for every inter-beat-interval was calculated. We analyzed the spectral densities of the resampled time series in the 0.07–0.14 Hz band as well as the 0.15–0.4 Hz band. The power in the low frequencies is related to sympathetic activity, whereas the high frequency band is almost completely influenced by the parasympathetic nervous system in addition to respiration [14]. The low band is therefore widely regarded as the better measure for workload and stress.

## 3 Results and Discussion

All data was submitted to a one-way repeated measures analysis of variance (ANOVA) for the factor of time delay.

The questionnaire data showed an effect of system delay. The overall workload follows a linear trend (F(3,21)=22.44, p<0.01), indicating that increases in system delay induced higher subjective workload in our participants. The same linear trend can be found for the independent scales of the NASA-TLX. Interestingly, even though the self-rated performance decreased and the frustration increased, the participants did not “give up” on the task but increased their effort accordingly. In addition, the overall composition of the workload stayed the same (see figure 2). This suggests that the experimental manipulation of system delay increased subjective workload without changing the nature of the task itself.

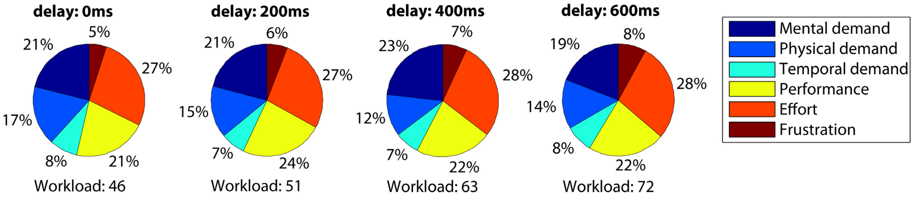

**Fig. 2.** Subjective workload increases with system delays but does not vary in its composition

As is the case with the self-rated performance, the objective error in compensating for the disturbance increased with increasing system delay (F(3,21)=12.65, p<0.01). Therefore, it follows that time delays have a deteriorating effect on the control task.

This deterioration in performance with increasing system delays evoked corrective inputs from the participants who tried to keep the vehicle stable. This is indicated by a linear increase in the power of the stick activity between 0.1 to 0.5 Hz (F(3,21)=36.32, p<0.01; see Figure 3). Since these disturbances take

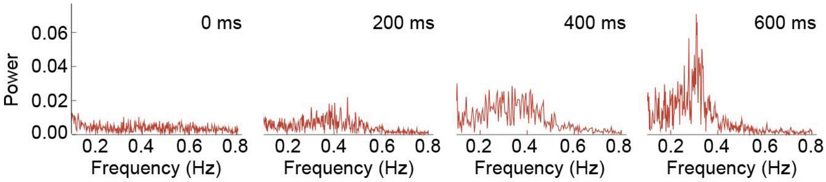

**Fig. 3.** Increasing input activity at frequencies higher than the disturbance. This can lead to pilot induced oscillations.

place at lower frequencies than these inputs, this behavior can destabilize the vehicle even more, resulting in pilot induced oscillations.

In correspondence with subjective workload measurements, an increase in system delays also resulted in higher skin conductance (F(3,21)=5.72, p=0.01, see figure 4). This indicates that participants experienced stress and arousal.

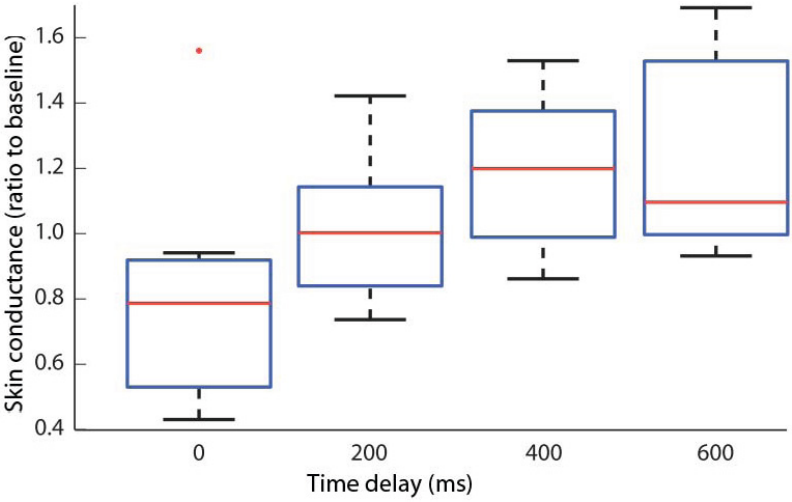

**Fig. 4.** The skin conductance increases linearly with increasing time delay

Nonetheless, the ECG measures for heart-rate and heart-rate variability did not show any significant changes to the manipulation of system delays. In addition, we did not find any changes between the test trials and breaks. Therefore, ECG based measures are not a reliable metric for stress and arousal in the current control task.

## 4 Conclusions

Overall, our findings show that delays between the control input and system response can impair control performance, elicit pilot-induced oscillations and increase workload, both in terms of self-reported and physiological measures. This is an important point to note in the design of virtual simulation systems, such as driving and flight simulators, that are intended for the purpose of training closed-loop control.

A system that is slow in responding to the human operator's input could induce the human operator to submit a larger response than is required for precise maneuvers. This results in a larger than intended vehicle response that needs to be corrected for subsequently. It could, thus, result in more errors than necessary and even instill counter-productive behavior that will have to be unlearned in the real world.

Our findings indicate that this loss of control has an impact on the operator's perceived and physiological workload. Therefore, system delays have a genuine influence on the operator's conscious sense of well-being as well as his physiological system.

In this work, we demonstrated that skin conductance activity can offer a complementary approach to the use of questionnaires. Changes in heart-based measurements might be too slow to indicate the changes in stress levels experienced by the human operator in our current experimental task. In contrast to a questionnaire, an unobtrusive measure such as this can be employed to assess multiple maneuvers in a complex mission. In addition to the traditional assessment of novel controller systems for their handling qualities, skin conductance measurements can allow the same systems to be evaluated for their physiological comfort.

To conclude, we demonstrate that system delays can detrimentally affect control performance due to pilot-induced oscillations. This has an adverse effect on the perceived workload of the operator as well as on his physiological system. The approach described here is a viable protocol for the evaluation of novel controller systems and simulators intended for closed-loop control.

## References

1. Allison, R., Harris, L., Jenkin, M., Jasiobedzka, U., Zacher, J.: Tolerance of temporal delay in virtual environments. In: Proceedings IEEE Virtual Reality 2001, pp. 247–254 (2001)
2. Middendorf, M., Lusk, S., Whitley, J.: Power spectral analysis to investigate the effects of simulator time delay on flight control activity. In: AIAA Flight Simulation Technologies Conference, pp. 46–52 (1990)
3. Wildzunas, R.M., Barron, T.L., Wiley, R.W.: Visual display delay effects on pilot performance. Aviation, Space, and Environmental Medicine 67, 214–221 (1996)
4. Jennings, S., Reid, L.D., Craig, G., Kruk, R.V.: Time Delays In Visually Coupled Systems During Flight Test and Simulation. Journal of Aircraft 41, 1327–1335 (2004)

5. Middendorf, M., Fiorita, A., McMillan, G.: The effects of simulator transport delay on performance, workload, and control activity during low-level flight. In: AIAA Flight Simulation Technologies Conference (1991)
6. Riccio, G., Cress, J., Johnson, W.: The effects of simulator delays on the acquisition of flight control skills: Control of heading and altitude. In: Proceedings of the Human Factors and Ergonomics Society Annual Meeting, pp. 1186–1290 (1987)
7. Zhao, C., Zhao, M., Liu, J., Zheng, C.: Electroencephalogram and electrocardiograph assessment of mental fatigue in a driving simulator. Accident; Analysis and Prevention 45, 83–90 (2012)
8. Chattopadhyay, P., Bond, A., Lader, M.: Characteristics of galvanic skin response in anxiety states. Journal of Psychiatric Research 12, 265–270 (1975)
9. Perfect, P., Jump, M., White, M.D.: Development of handling qualities requirements for a personal aerial vehicle. In: Proceedings of the 38th European Rotorcraft Forum, Amsterdam, Netherlands (2012)
10. Perry, A.R.: The Flightgear flight simulator. In: 2004 USENIX Annual Technical Conference, Boston, MA (2004)
11. Hart, S., Staveland, L.: NASA Task Load Index (TLX) v1. 0 users manual (1986)
12. Montagu, J., Coles, E.: Mechanism and measurement of the galvanic skin response. Psychological Bulletin 65, 261–279 (1966)
13. Tattersall, A., Hockey, G.: Level of operator control and changes in heart rate variability during simulated flight maintenance. The Journal of the Human Factors 37, 682–698 (1995)
14. Camm, A., Malik, M.: Heart Rate Variability: Standards of Measurement. European Heart Journal of the Physiological Interpretation and Clinical Use (1996)