Between-Simulator Comparison of Motion Filter Order and Break Frequency Interaction Effects on Pilot Control Behavior

Daan M. Pool*
Delft University of Technology, Delft, Zuid Holland, 2600 HS, Netherlands

Peter M. T. Zaal†
Metis Technology Solutions, Inc., NASA Ames Research Center, Moffett Field, CA, 94035

Marc A. Pieters‡
San José State University, NASA Ames Research Center, Moffett Field, CA, 94035
Delft University of Technology, Delft, Zuid Holland, 2600 HS, Netherlands

Olaf Stroosma§ and Max Mulder¶
Delft University of Technology, Delft, Zuid-Holland, 2600 HS, Netherlands

This paper investigates the interaction effects of motion filter order and break frequency on pilots' manual control behavior and control performance using two simulators. Eighteen pilots performed the experiment in the Vertical Motion Simulator (VMS) at NASA Ames Research Center and twenty pilots in the SIMONA Research Simulator at Delft University of Technology. The experiment used a full-factorial design with three motion filter orders (first-, second-, and third-order) and two filter break frequencies (0.5 and 2.0 rad/s), in addition to reference no-motion and full-motion conditions. Key task variables, such as the quality of the motion and visual cues and the characteristics of the sidestick, were matched across both simulators. Overall, the expected effects of filter order and break frequency variations were found, with both increasing order and increasing break frequency causing pilots to use less motion feedback in their control strategy, resulting in lower tracking performance. Furthermore, across the wide range of filter orders tested in the experiment, the existing Sinacori-Schroeder motion fidelity criterion was found to be a good predictor of the interaction effects of both filter settings on pilot control behavior. For the same motion condition, there was a consistent offset in the results between simulators, due to the more high-gain control strategy adopted by a number of

*Assistant Professor, Control and Simulation Division, Faculty of Aerospace Engineering, P.O. Box 5058, 2600 GB Delft, Netherlands; d.m.pool@tudelft.nl. Senior Member AIAA.

†Principal Aerospace Engineer, SimLabs, NASA Ames Research Center, Moffett Field, CA, 94035; peter.m.t.zaal@nasa.gov. Associate Fellow AIAA.

‡Research Scholar, Human Systems Integration Division, NASA Ames Research Center, Moffett Field, CA, 94035, and MSc. Student, Control and Simulation Division, Faculty of Aerospace Engineering, Delft University of Technology, Delft, Netherlands; marc.a.pieters@gmail.com. Student Member AIAA.

§Researcher, Control and Simulation Division, Faculty of Aerospace Engineering, P.O. Box 5058, 2600 GB Delft, Netherlands; o.stroosma@tudelft.nl. Senior Member AIAA.

¶Professor, Control and Simulation Division, Faculty of Aerospace Engineering, P.O. Box 5058, 2600 GB Delft, Netherlands; m.mulder@tudelft.nl. Associate Fellow AIAA.

the VMS pilots. Still, the observed relative trends in pilot control behavior and performance between motion conditions were equivalent in both simulators and thus accurately replicated.

Nomenclature

e	=	error signal, deg
f_d	=	disturbance forcing function, deg
f_t	=	target forcing function, deg
H_c	=	controlled dynamics
H_{mf}	=	motion filter
H_{mot}	=	motion hardware dynamics
H_{shp}	=	motion shaping filter
H_{stk}	=	stick dynamics
H_{SRS}	=	SRS motion dynamics
$H_{p_{mot}}$	=	pilot motion response
$H_{p_{vis}}$	=	pilot visual response
H_{VMS}	=	VMS motion dynamics
K_m	=	pilot motion gain, –
K_{mf}	=	motion filter gain, –
K_S	=	gain of motion filter at 1 rad/s, –
K_v	=	pilot visual gain, –
n	=	pilot remnant, deg
O_{mf}	=	motion filter order, –
s	=	Laplace operator, rad/s
t	=	time, s
T_L	=	pilot lead time constant, s
T_m	=	measurement time, s
u	=	pilot control input, deg
u_m	=	motion contribution to control input, deg
u_v	=	visual contribution to control input, deg
δ_e	=	elevator deflection, deg
ζ_{nm}	=	neuromuscular damping, –
θ	=	pitch angle, deg

μ = average, –

σ = standard deviation, –

τ_m = motion time delay, s

τ_v = visual time delay, s

φ_m = open-loop phase margin, deg

Φ_S = phase of motion filter at 1 rad/s, deg

ω_c = open-loop crossover frequency, rad/s

ω_{mf} = motion filter break frequency, rad/s

ω_n = stick natural frequency, rad/s

ω_{nm} = neuromuscular frequency, rad/s

ω_{phg} = phugoid frequency, rad/s

ω_{sp} = short period frequency, rad/s

Abbreviations

ERP = eye reference point

ICR = instantaneous center of rotation

IDMS = Image Delay Measurement System

OMCT = Objective Motion Cueing Test

PFD = primary flight display

RMS = root mean square

SRS = SIMONA Research Simulator

VAF = variance accounted for

VDMS = Visual Delay Measurement System

VMS = Vertical Motion Simulator

I. Introduction

Despite a lack of consensus on the universal benefit of simulator motion for pilot training [1–7], there is still a strong emphasis on training in the presence of motion feedback, as demonstrated by recent advances in motion fidelity standards for FSTDs [8] and the new requirement for airline pilots to receive stall recovery training in moving-base flight simulators [9–11]. To keep simulators within the constraints of their motion workspaces, motion cueing algorithms use washout filters to attenuate the true motion simulated by the aircraft dynamic model [12–14]. The characteristics of washout filters in every simulator degree-of-freedom are defined by several tunable parameters, such as gains, break

frequencies, and damping coefficients. Even though many practical cueing algorithms contain some form of non-linear adaptation over the available motion space, this study, like many others, only considers constant parameters.

Numerous studies have investigated the influence of these washout-filter parameters on pilot control behavior and performance [5, 10, 15–27] in an effort to provide guidelines for washout filter tuning. The effect of the washout filter *order*, has received surprisingly little attention, while, as documented in literature overviews compiled in [15] and [17], it does vary between different previous motion cueing experiments. Detailed investigation of the effects of filter order on control behavior and performance is essential for adding to our knowledge of how pilots are affected by motion cueing, as the selected filter order also potentially interacts with the effects of other motion filter parameters, especially those that affect phase discrepancies (e.g., break frequencies).

Despite the importance of verifying experimental results by other researchers, replicating simulator experiments in the field of human perception and cueing, as in many other fields, has been problematic and only sparsely attempted. Correia-Graćio et al. [28] investigated visual-vestibular coherence zones using two simulators. Grant et al. [29, 30] and Ellerbroek et al. [31] both replicated an experiment performed by Schroeder [32] with a helicopter yaw-capture task under varying sway and yaw motion conditions. Jex et al. [33, 34] replicated their own roll-tracking experiment on two simulators and attempted to find the source of differences in the effects of simulator motion on manual control behavior between the two simulators. In these sets of experiment replications, mostly matching results were found between replicated experiments. However, even though considerable effort was spent on matching the experimental setups (see, e.g., [31]), some critical differences in experiment outcomes and conclusions were also reported. Hence, as also stated in [3], to be able truly generalize the main findings regarding the important and subtle effects of simulator motion cueing on pilots' control behavior, and to develop verifiable standards for simulator cueing, more replications of pilot-in-the-loop simulator experiments are needed.

This paper presents the results of an experiment with thirty-eight general aviation pilots that was performed to investigate, and compare, the effects of motion-filter order and break frequency on manual control behavior and performance and to verify these effects across simulators. Pilot control behavior was measured in a compensatory pitch-tracking task, based on an earlier experiment [35], which was performed under eight different motion conditions by all participants. The experiment was conducted on two simulators – the Vertical Motion Simulator (VMS) at NASA Ames and the SIMONA Research Simulator (SRS) at TU Delft – to compare the measured effects of motion cueing variations across the replicated experiments.

The authors aim to make the following contributions to literature. First, this is the first study focusing entirely on the interaction effects of motion-filter order and motion filter break-frequency settings on manual control behavior and performance. Second, the experiment uses two simulators with different motion capabilities to directly verify the generalizability of the results across different platforms. Third, the setups in both simulators are documented in full detail to support future verification experiments. Finally, with a large group of general aviation pilots as participants

on both simulators (i.e., 18 for the VMS and 20 for the SRS), the study uses a mixed-effects model approach [36] to perform statistical analysis on both the measured within- and between-experiment effects and their interactions.

The paper is structured as follows. Sections II and III present the task and experiment design, respectively. Results are presented in Section IV and discussed in Section V. Section VI provides the conclusions.

II. Control Task

Fig. 1 illustrates the manual pitch control task that participants performed in both the VMS and SRS. The pitch attitude tracking task, with a compensatory display, was based on a previous experiment by Zaal and Zavala [35]. Participants minimized the pitch error e, representing the difference with the aircraft pitch angle θ and the reference pitch angle f_t ($e = f_t - \theta$) by making control inputs u using a sidestick with dynamics $H_{stk}(s)$. The display represented a simplified primary flight display (PFD) on which the vertical displacement of the horizon line indicated the current tracking error e, see Fig. 1 and Appendix A.D for details. The control inputs δ_e acted on the aircraft pitch dynamics $H_\theta(s)$, which resulted in pitch angle θ.

Pilots perceived the motion cues resulting from a cascaded motion filter, with dynamics $H_{mf}(s)$, and motion system $H_{mot}(s)$. In the VMS, the motion system was characterized by the motion system dynamics, $H_{mot}(s) = H_{VMS}(s)$. In the SRS a motion shaping filter $H_{shp}(s)$ was used in conjunction with the SRS motion system dynamics, $H_{mot}(s) = H_{shp}(s) \cdot H_{SRS}(s)$, to match the motion system response of the SRS to the VMS (see Appendix A.B). Applying both a target forcing function f_t and disturbance forcing function f_d allows for the identification of a multi-channel quasi-linear human pilot model consisting of a visual response function $H_{p_{vis}}(s)$, a motion response function $H_{p_{mot}}(s)$, and a remnant signal n [37, 38]. Note that as pilot and stick dynamics cannot be separated for our tracking task, the identified visual and motion response dynamics also include the stick dynamics shown in Fig. 1, i.e., $H_{p_{vis}}(s) = H^*_{p_{vis}}(s)H_{stk}(s)$ and $H_{p_{mot}}(s) = H^*_{p_{mot}}(s)H_{stk}(s)$.

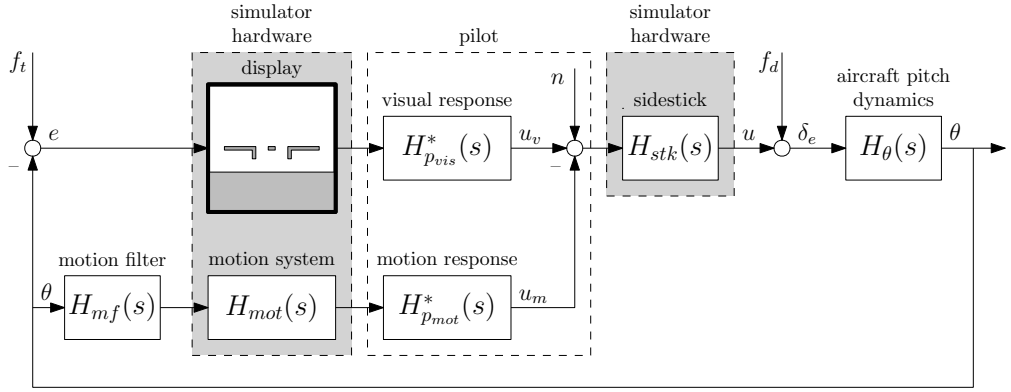

Fig. 1 Block diagram of the pitch control task and the human pilot model.

A. Controlled Dynamics

The aircraft pitch dynamics, $H_\theta(s)$, were the same as those used in the experiment of [35]:

$$H_\theta(s) = \frac{\theta(s)}{\delta_e(s)} = \frac{28.4474 \cdot \left(346.5s^2 + 32.03s + 1\right)}{\left(245.6s^2 - 3.409s + 1\right) \cdot \left(2.105s^2 + 0.9387s + 1\right)} \tag{1}$$

These represent the pitch dynamics of a mid-size twin-engine commercial transport aircraft with a weight of 185,000 lbs, trimmed close to its stall point, at 41,000 ft with an indicated airspeed of 150 kts. The dynamics feature a stable short period eigenmode (ω_{sp} = 0.6892 rad/s) and an unstable phugoid eigenmode (ω_{phg} = 0.0638 rad/s), with eigenvalues at $\lambda_{1,2} = -0.2230 \pm 0.6522i$ and $\lambda_{3,4} = 0.0069 \pm 0.0634i$ in the complex plane, respectively. These dynamics required pilots to perform high-frequency lead equalization ($T_L \approx 1/0.6892$ rad/s = 1.4509 s). This implies that for these aircraft pitch dynamics a considerable benefit of motion feedback is expected [39] and that the effects measured in the experiment are thus representative for dynamics with similar compensation demands.

In a pitch control task with a fixed-wing aircraft, a pilot feels the combination of vertical (heave) motion of the aircraft's instantaneous center of rotation (ICR) and the vertical accelerations resulting from the pitch rotation and the offset between the pilot station and the ICR (pitch heave), in addition to the rotational pitch motion [35, 40]. To ensure the task could be performed with 1-to-1 motion feedback within the motion space of both simulators, only the rotational pitch motion and the pitch heave with respect to the ICR were simulated. A previous study showed that not simulating ICR heave in a pitch control task did not significantly affect pilot control behavior [40]. The ICR pitch heave response to pitch variations is defined by:

$$H_{a_{z_{\theta,ICR}}}(s) = \frac{a_{z_{\theta,ICR}}(s)}{\theta(s)} = -11.49s^2 \tag{2}$$

Eq. (2) shows that the pilot station was located 11.49 m in front of the ICR. Analogous to ICR pitch heave, a vertical offset between the pilot station and the ICR results in ICR "pitch surge". In the considered aircraft, this vertical offset was not present, such that no pitch surge was present. Any surge movement of the ICR was also neglected.

B. Motion Filter

In this experiment, only aircraft pitch and heave motion was replicated using the simulators' motion systems. Furthermore, to ensure congruent pitch and heave motion cues, identical linear motion filters $H_{mf}(s)$ were always used in the pitch and heave axes. The experiment manipulated three high-pass motion filter orders as typically used in motion cueing algorithms [15, 17]: first-, second-, and third-order high-pass filters. The transfer functions for these different motion filters are given by Eq. (3), where K_{mf} is the motion filter gain, ζ_{mf} the damping ratio, and ω_{mf} the

break frequency:

$$H_{mf}(s) = K_{mf} \cdot \underbrace{\overbrace{\frac{s}{s+\omega_{mf}}}^{\text{first order}} \cdot \overbrace{\frac{s^2}{s^2+2\zeta_{mf}\omega_{mf}s+\omega_{mf}^2}}^{\text{second order}}}_{\text{third order}} \tag{3}$$

Note from Eq. (3) that, for example, for a second-order filter, only the transfer function term labeled "second order" would be used. Furthermore, it should be noted that unlike the independent low break frequency typically used for the first-order filter term in third-order motion filters [13, 25], here the same ω_{mf} was used deliberately for all filter break frequencies to explicitly measure the interaction effects of break frequency and filter order. The filters varied between the tested experimental conditions, see Section III.A.

C. Pilot Model

To explicitly quantify the adopted pilot manual control behavior, linear frequency response functions were identified for both pilots' visual and the motion channels, as depicted in Fig. 1. McRuer and Jex [41] have shown that, with compensatory displays, pilots adapt to the controlled dynamics such that the open-loop response approximates integrator dynamics in the crossover frequency region. For the controlled dynamics of Eq. (1), pilots thus needed to generate lead in the region of the crossover frequency. Hence, the same models for pilots' visual and motion responses as also considered in [35] were used here, which are defined in Eq. (4) and Eq. (5), respectively:

$$H_{p_{vis}}(s) = H^*_{p_{vis}}(s)H_{stk}(s) = K_v\,(1+T_L s)\,e^{-\tau_v s}\frac{\omega_{nm}^2}{s^2+2\zeta_{nm}\omega_{nm}s+\omega_{nm}^2} \tag{4}$$

$$H_{p_{mot}}(s) = H^*_{p_{mot}}(s)H_{stk}(s) = sK_m e^{-\tau_m s}\frac{\omega_{nm}^2}{s^2+2\zeta_{nm}\omega_{nm}s+\omega_{nm}^2} \tag{5}$$

As can be verified from Eqs. (4) and (5), the pilot model had a total of seven parameters that quantified pilots' control behavior. Pilot equalization, or adaptation to the controlled dynamics, is captured with the equalization parameters: the visual gain K_v, the motion gain K_m, and the lead time constant T_L. Pilot limitations are captured with the visual perception time delay τ_v, the motion perception time delay τ_m, and neuromuscular dynamics represented by a second-order filer with damping ζ_{nm} and break frequency ω_{nm} [42].

D. Forcing Functions

Two forcing functions were used in the pitch tracking task, a target and a disturbance signal, which made the task a combined target-following and disturbance-rejection task. Using two independent forcing function signals allows for the identification of $H_{p_{vis}}$ and $H_{p_{mot}}$ of the pilot model as introduced in Section II.C [37, 38]. Both forcing functions, which exactly matched those used in [35], were defined as sum-of-sines signals:

$$f_{t,d}(t) = \sum_{k=1}^{N_{t,d}} A_{t,d}(k) \sin\left[\omega_{t,d}(k)t + \phi_{t,d}(k)\right], \tag{6}$$

where $A_{t,d}(k)$, $\omega_{t,d}(k)$ and $\phi_{t,d}(k)$ are the amplitude, frequency and phase of the k^{th} sine in the target and disturbance forcing functions f_t and f_d, respectively. Each forcing function consisted of $N_{t,d} = 10$ sinusoids. Table 1 provides the parameter values for both the f_t and f_d signals. The frequencies $\omega_{t,d}$ for the sinusoids were all integer multiples $n_{t,d}$ of the measurement time base frequency, $\omega_m = 2\pi/T_m = 2\pi/81.92$ s = 0.0767 rad/s, to avoid spectral leakage. The integer multiples were selected to ensure that the typical frequency range of human control (0.1-15 rad/s) was covered with regular intervals on a logarithmic scale [35]. Both the target forcing function f_t and the disturbance forcing function f_d had a variance of 0.4 deg^2.

Experiment runs lasted 94.92 seconds. The first three seconds contained no forcing functions. In the following five seconds the forcing functions were gradually faded in, to allow pilots to stabilize the controlled element. Next, a measurement window of 81.92 seconds was used for the analysis. The last 5 seconds were a ramp down of the forcing functions in order to return the simulator to its initial position gradually.

Table 1 Forcing function properties [35].

Target, f_t				Disturbance, f_d			
n_t, –	ω_t, rad/s	A_t, deg	ϕ_t, rad	n_d, –	ω_d, rad/s	A_d, deg	ϕ_d, rad
3	0.2301	0.5818	-1.4796	2	0.1534	0.0105	0.1355
6	0.4602	0.5306	-0.0745	5	0.3835	0.0098	-0.1664
13	0.9971	0.3711	0.7006	11	0.8437	0.0091	2.9016
27	2.0709	0.1674	-1.9563	23	1.7641	0.0283	5.6383
41	3.1447	0.0901	-2.8131	37	2.8379	0.0403	2.8648
53	4.0650	0.0605	2.1026	51	3.9117	0.0477	4.8718
73	5.5990	0.0375	-2.6178	71	5.4456	0.0569	1.0245
103	7.9000	0.0238	2.2550	101	7.7466	0.0725	5.0337
139	10.6612	0.0174	-0.6739	137	10.5078	0.0967	4.1487
194	14.8796	0.0135	0.1942	191	14.6495	0.1458	0.4274

III. Experiment

A. Independent Variables

The experiment tested one between-subject variable and two within-subject independent variables. The between-subject independent variable had two levels, VMS and SRS, and accounts for possible differences found between the two simulators used. The within-subjects variables were: (1) motion-filter order O_{mf}, with three levels (1st, 2nd, and 3rd order) and motion-filter break frequency ω_{mf}, with two levels (0.5 and 2.0 rad/s). The filter break frequency was selected as a second independent variable to act as a possible reference for the size of the effects due to O_{mf} variations, as well as for studying possible interaction effects, that is, to show whether effects of motion filter order depend on ω_{mf}, or not. A full-factorial design based on these two independent variables was considered, resulting in six experimental conditions (C1-C6), listed in Table 2.

In addition, two motion conditions were added as a reference, the no-motion condition (fixed base, FB) and full motion condition (FM). The no-motion condition FB was included to isolate the effects of the simulator motion systems. The full-motion condition was used as a baseline for the effects of the motion filter variations in conditions C1-C6. For condition FM a second-order filter with a very low break frequency of $\omega_{mf} = 0.2$ rad/s was used to prevent simulator drift, as also done in earlier VMS experiments [5, 35]. In all conditions with simulator motion, the motion filter gain was set to $K_{mf} = 1.0$, and the damping ratio was set to $\zeta_{mf} = 1/2\sqrt{2} = 0.707$. Fig. 2 shows the motion filter settings for conditions C1-C6 and FM in a Bode plot. As is clear from Fig. 2, the full-motion condition only resulted in minor motion attenuation at low frequencies, well below the 2.5-5 rad/s frequency range where the pilot-vehicle system crossover frequency is expected [43] and frequencies >3 rad/s where motion feedback is most beneficial for improving task performance [39]. Furthermore, Fig. 2 shows that both increasing filter order and ω_{mf} results in significantly increased attenuation and phase lead, also in these task-relevant frequency ranges.

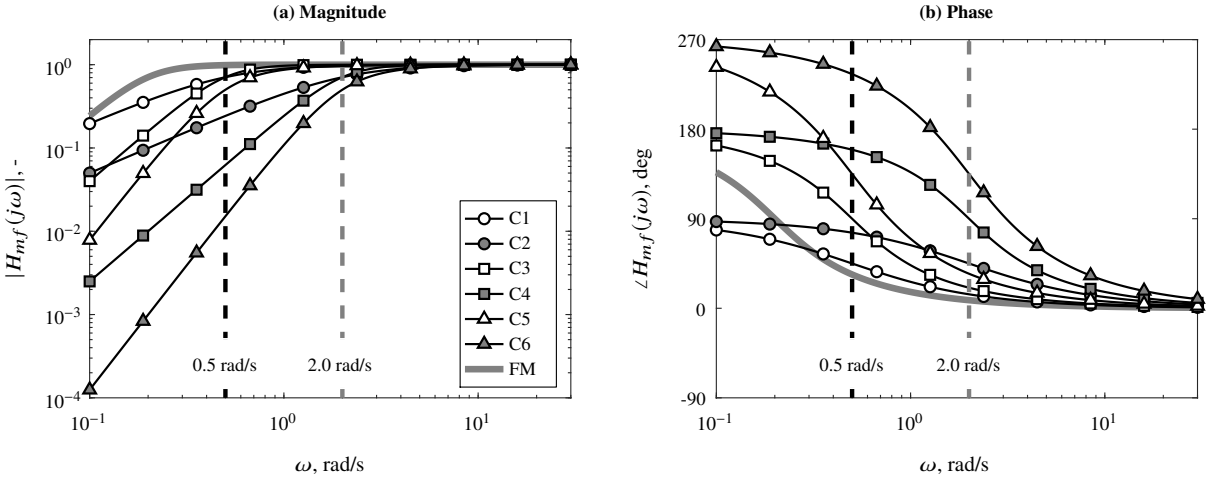

Fig. 2 Frequency responses of the motion filters used for conditions C1-C6 and FM.

Table 2 Experimental conditions.

Condition	Filter Order O_{mf}, –	Break Frequency ω_{mf}, rad/s
FM	2	0.2
C1	1	0.5
C2	1	2.0
C3	2	0.5
C4	2	2.0
C5	3	0.5
C6	3	2.0
FB	–	–

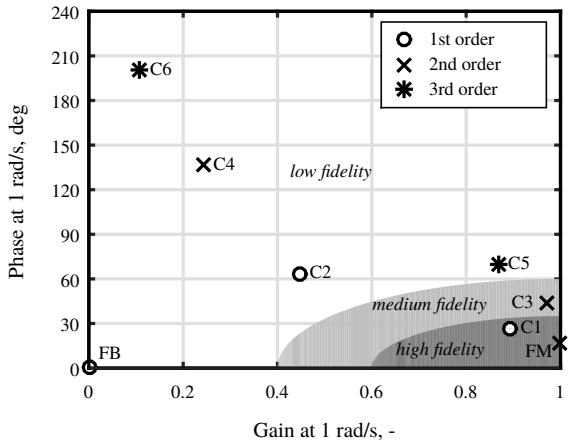

Fig. 3 Sinacori-Schroeder plot with motion conditions.

Fig. 3 shows the predicted fidelity of the eight motion conditions against the fidelity criteria proposed by Sinacori

[44], with updated limits from Schroeder [32]. The Sinacori-Schroeder plot shows the gain K_S and phase Φ_S of the motion filter at 1 rad/s with respect to high-, medium- and low-fidelity boundaries originally defined based on experiments with second-order motion filters [32]. With increasing filter order O_{mf}, the conditions move further away from the high fidelity region. This effect is stronger with the higher motion filter break frequency $\omega_{mf} = 2.0$ rad/s, which illustrates the interaction effect between break frequency and order, and thus the importance of using this parameter as an independent variable.

B. Apparatus

The experiment was first conducted in the VMS at NASA Ames Research Center using the transport cab (T-CAB), see Fig. 4(a). The SRS at Delft University of Technology (Fig. 4(b)) was then used to replicate the VMS experiment. The simulator cockpits are shown in Figs. 4(c) and (d), respectively. Participants were seated in the right-hand seat, with the PFD in front and the sidestick to their right-hand side, see Figs. 4(c) and (d). The out-the-window visual system was not used. Participants wore noise-cancelling headphones with additional brown noise resembling aircraft engines played over the headphones, which successfully masked any sounds made by the motion system actuators. More details on the different simulator hardware systems and the equalization between simulators are provided in Appendix A.

C. Participants

Eighteen pilots participated in the VMS experiment and 20 pilots participated in the SRS experiment. All were active general aviation pilots. Table 3 presents information on the pilot population. Four VMS pilots had considerably more flight hours than the rest: 5,300, 2,800, 1,637 and 1,200 hours. Similarly, two SRS pilots had flown considerably more hours than the rest: 6,800 and 1,018 hours. Most pilots in both groups had prior experience in fixed-base or full-motion training simulators. Most of the VMS pilots also had participated in earlier simulator experiments in the VMS (for example, [35] and [5]), whereas the recruited SRS pilots generally had no prior experience from earlier experiments.

Table 3 Overview of participating pilots' relevant characteristics.

	Age, years		Flight Hours		Simulator Hours	
	μ	σ	μ	σ	μ	σ
VMS	28.9	4.97	751	1341	45.7	91.0
SRS	31.5	5.52	636	1455	27.9	60.2

D. Procedures

At the start of the experiment, pilots were briefed by explaining the purpose of the experiment, the task, and the experiment procedures. Pilots were instructed to focus on continuous minimization of the tracking error e, by

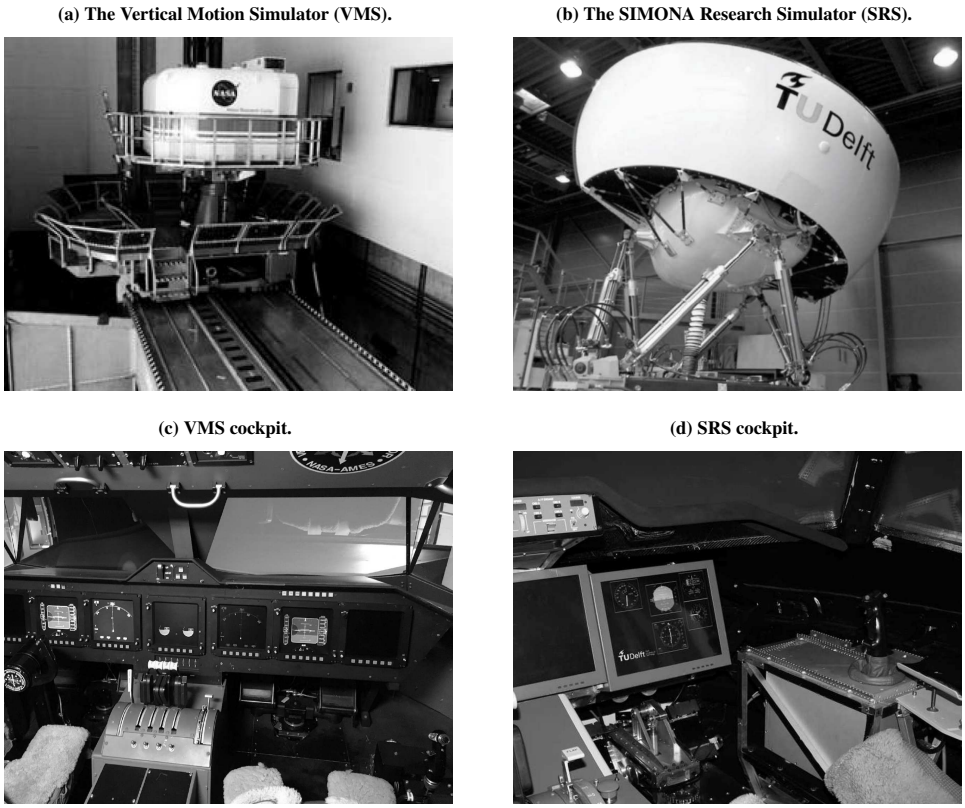

Fig. 4 The Vertical Motion Simulator (VMS) and the SIMONA Research Simulator (SRS) and their respective flight decks.

keeping the aircraft symbol as accurately as possible on the horizon of the primary flight display (see Fig. 1), and to provide smooth and continuous control inputs. Furthermore, the fact that in the pitch tracking task tracking errors were induced by both the target and disturbance signals was explained to all pilots. No specifics about the (number of) motion conditions were given, except that a no-motion condition was present. As a reference for their own level of task performance, pilots were informed of the best performance score in terms of root mean square (RMS) of the error signal of all participants so far and encouraged to improve it. After each run the PFD showed the run score to give the pilots unbiased and consistent feedback on their performance.

The experiment consisted of three simulator sessions, all performed on the same day, with breaks in between sessions. The participants performed 24, 20 and 12 tracking runs in these three consecutive sessions, respectively. The participants were informed that they could take additional or longer breaks if desired (for example, when fatigued). Throughout the experiment, the eight different conditions were tested in mixed and randomized order, i.e., in sets of eight runs, one run of each condition, with a different randomized conditions order for each set. The first 16 runs (two sets of all conditions) were used as initial training and the last 40 runs (five sets of all conditions) were used as the measured dataset for results analysis. Every pilot thus performed each experimental condition seven times in total.

E. Dependent Measures

The experiment goal was to investigate if and how the motion washout filter order O_{mf} and the break frequency ω_{mf} influenced pilot control behavior and performance, and whether observed trends were comparable between two flight simulators. Pilot control behavior and performance metrics were the dependent measures of interest. The root mean square (RMS) values of the error signal (RMS_e) and the control signal (RMS_u) were used as the measures of performance and control activity, respectively. A lower RMS_e indicates a lower overall error score and hence improved performance. A higher RMS_u indicates increased control activity by the pilot.

In addition to these performance metrics, the seven parameters of the pilot model defined in Eqs. (4) and (5) – i.e., K_v, K_m, T_L, τ_v, τ_m, ζ_{nm} and ω_{nm} – were estimated using a time-domain multi-channel pilot model parameter estimation technique [38]. This technique uses a genetic algorithm to provide an initial estimate for the parameters, which is subsequently refined by a gradient-based Gauss-Newton estimation routine. Non-parametric frequency response estimates of $H_{p_v}(j\omega)$ and $H_{p_m}(j\omega)$ (not shown here for brevity) were used to verify all time-domain pilot model estimates. The variance accounted for (VAF), which expresses how much of the control signal u can be explained by the fitted pilot model in a percentage, was determined as a measure of model accuracy. Furthermore, using the estimated pilot model transfer functions, the variances of the control signal contributions from both the visual and motion channels (see Fig. 1), $\sigma^2_{u_v}$ and $\sigma^2_{u_m}$, respectively, were calculated. The fraction of these variances $\sigma^2_{u_m}/\sigma^2_{u_v}$ defines the relative contribution of the motion channel to the pilot's control inputs [17].

Finally, the crossover frequencies and phase margins of the open-loop dynamics were used to quantify pilot performance in the frequency domain in attenuating the target and disturbance signals [40]. Using Fig. 1, an open-loop response can be constructed for both the target and disturbance inputs, as defined in Eqs. (7) and (8). From estimated non-parametric open-loop frequency response estimates according to the middle terms in Eqs. (7) and (8), the open-loop crossover frequencies and phase margins for both the target and disturbance loops – i.e., $\omega_{c,t}$, $\omega_{c,d}$, $\varphi_{m,t}$ and $\varphi_{m,d}$ – were estimated.

$$H_{ol,t}(s) = \frac{\theta(s)}{E(s)} = \frac{H_{p_{vis}}(s)H_\theta(s)}{1 + H_{mf}(s)H_{mot}(s)H_{p_{mot}}(s)H_\theta(s)} \tag{7}$$

$$H_{ol,d}(s) = -\frac{U(s)}{\delta_e(s)} = H_\theta(s)\left[H_{p_{vis}}(s) + H_{mf}(s)H_{mot}(s)H_{p_{mot}}(s)\right] \tag{8}$$

F. Data Analysis

One pilot from each simulator group was omitted from the analysis. The VMS pilot did not perform the tracking task at a sufficiently accurate and consistent level of performance to enable meaningful data analysis, while the SRS pilot was unable to complete the experiment. Hence, data from 17 pilots for the VMS and 19 pilots for the SRS experiments were included in the analysis.

Linear mixed-effects models [36] were used to determine whether the independent variables of the experiment introduced significant effects in the dependent measures across experiment conditions. Two separate mixed-effect model analyses were performed for each dependent measure. Using only the data from conditions C1-C6 (and hence omitting the data for the reference FB and FM conditions), the first analysis tested for effects of simulator (VMS or SRS), motion filter order O_{mf} (first-, second-, or third-order), and motion filter break frequency ω_{mf} (0.5 or 2.0 rad/s). The mixed-effects model used for this analysis included "Simulator", O_{mf}, ω_{mf}, and all their interactions as fixed effects. The different pilots were included as the random effect. The second analysis used the data from all experiment conditions to test for effects of simulator (VMS or SRS) and motion condition (C1-C6, FB, and FM). In this model, "Simulator" and "Motion condition", and the interaction between both factors were used as the fixed effects, while again pilot was used as the random effect. The second mixed-effects model was used to verify any effects of simulator, found with the main first analysis, but now including the reference motion conditions FB and FM. The results of this analysis are presented in Appendix B. In both mixed-effects models only random intercepts were considered, i.e., no random slopes were introduced.

Before fitting the mixed-effects models, two major outliers due to innaccurate pilot model estimates were removed for $\sigma^2_{u_m}/\sigma^2_{u_v}$ (i.e, $\sigma^2_{u_m}/\sigma^2_{u_v} > 100\%$) and one for τ_m (i.e., $\tau_m > 1$ s). Assumptions of linearity, homoskedasticity, and normality of residuals were checked visually using residual plots and Q-Q plots for each model. For RMS_e, RMS_u, $\sigma^2_{u_m}/\sigma^2_{u_v}$, K_v, T_L, τ_m, and ζ_{nm} the model assumption of homoskedasticity was not met due to the fact that these measures were highly positively skewed. These dependent measures were transformed to make them more normally distributed using a Box-Cox transformation [45]. The lambda values for the transformations were -0.75, -0.06, 0.33, -0.4, -0.13, 0.22, and 0.03 for RMS_e, RMS_u, $\sigma^2_{u_v}/\sigma^2_{u_m}$, K_v, T_L, τ_m, and ζ_{nm}, respectively. All the dependent measure data met the assumption of homoskedasticity after the transformations were applied. No other significant violations of the assumptions for using mixed-effects models were present.

To facilitate statistical analysis of the effects of "Simulator", O_{mf}, ω_{mf} and their interactions, the mixed-effects models were progressively built up by adding the different fixed effects and interactions one-by-one, starting with the intercept-only model. Likelihood ratio tests between the models with and without a fixed effect determined the significance of adding that effect. The fixed-effects were added to the models in the order from left to right in Tables 4 to 7 and Table 11. Pairwise comparisons using a Tukey correction were performed as post-hoc tests to reveal significant differences between specific condition pairs when significant main effects or interactions were found by the mixed-effects models.

G. Hypotheses

Based on the expected fidelity of each motion condition against available motion fidelity criteria for high-pass motion filters [32, 44], see Fig. 3, and earlier experimental findings [3, 15, 16], two hypotheses were formulated

regarding the anticipated changes in the dependent measures due to a change in filter order O_{mf} or filter break frequency ω_{mf}:

H1: *Higher motion filter order O_{mf} reduces pilots' use of motion feedback and decreases task performance. These effects are stronger for higher values of ω_{mf}.* Increasing the motion filter order decreases the fidelity of the cued motion, see Fig. 3. This effect is especially strong for the higher motion filter break frequency $\omega_{mf} = 2.0$ rad/s. With increasing motion filter order more of the low frequency content of the aircraft pitch output is filtered out by the motion filter. In addition, a high O_{mf} introduces increased phase lead in the simulator motion. As these effects were reported in earlier experiments as a result of similarly degraded motion feedback [15–17], it was hypothesized that increased O_{mf} leads to a decrease in tracking performance (higher RMS_e), a decrease in visual gain K_v, an increase in the visual lead time-constant T_L, and an overall reduced contribution of motion feedback to pilots' control behavior (reduced $\sigma^2_{u_m}/\sigma^2_{u_v}$). Furthermore, lower disturbance crossover frequencies $\omega_{c,d}$ and higher disturbance phase margins $\varphi_{m,d}$ are expected.

H2: *Higher motion filter break frequencies ω_{mf} also decrease task performance and reduce pilots' use of motion feedback, and between the two tested ω_{mf} settings of 0.5 and 2.0 rad/s these effects are stronger than those that occur between the tested motion filter orders.* As shown in Fig. 2 and 3, increasing the motion filter break frequency ω_{mf}, for which more of the low-frequency content of the aircraft output will be filtered out by the motion filter, results in reduced motion fidelity, equivalent to the hypothesized effects of O_{mf}. Hence, when increasing ω_{mf} from 0.5 to 2.0 rad/s we expect to find the same effects on the dependent variables mentioned for Hypothesis H1, albeit stronger.

In addition to these hypotheses related to the effects of our independent variables (O_{mf} and ω_{mf}), we also formulated two hypotheses for the expected between-simulator differences, as the experiment was performed on two simulators. Considerable effort was put into equalizing the motion, visual, and control inceptor systems between simulators. Furthermore, the task and experimental conditions were the same and the participant group in each simulator was similar (i.e., aircraft types, experience, age). Hence, the following two hypotheses were formulated:

H3: *No significant differences in absolute value will be found in the considered dependent measures between both simulators.* Due to the efforts put into equalizing the VMS and SRS experiments, yielding identical control tasks in the two simulators, it was hypothesized that both simulators would deliver the same absolute results. In statistical terms, this means no statistically significant main effects of our between-group "Simulator" variable for our dependent measures are expected.

H4: *No significant differences in the relative trends in our dependent measures across experiment conditions will be found.* As described in Hypotheses H1 and H2, the applied variation in O_{mf} and ω_{mf} is expected to result in trends in key dependent measures that are known to capture pilots' adaptation to changes in motion filter settings. Equivalent to our Hypothesis H3, it is expected that the pilot control behavior data collected in both

simulators will show the same relative trends between the different motion conditions. In statistical terms, this means no significant interaction effects between our within-subjects independent variables (motion cueing variations, i.e., O_{mf} and ω_{mf}) and our between-group variable (i.e., "Simulator") are expected.

IV. Results

This section presents the combined results for the tracking performance, control activity, pilot model parameters and open-loop parameters. All dependent measures will be shown using boxplots. In each boxplot, a horizontal line indicates the median, and a circle indicates the mean. The bottom and top edges of the box indicate the first and third quartiles, respectively. The length of the upper whisker is the largest value that is no greater than the third quartile plus 1.5 times the interquartile range. The length of the lower whisker is the lowest value that is not smaller than the first quartile minus 1.5 times the interquartile range. Points outside of the whiskers are outliers and marked by plus symbols. All boxplots depict the data corrected for between-subject variability. Below each figure, a table is included with the corresponding likelihood ratio test results for the (first) mixed-effect model analysis, as described in Section III.F.

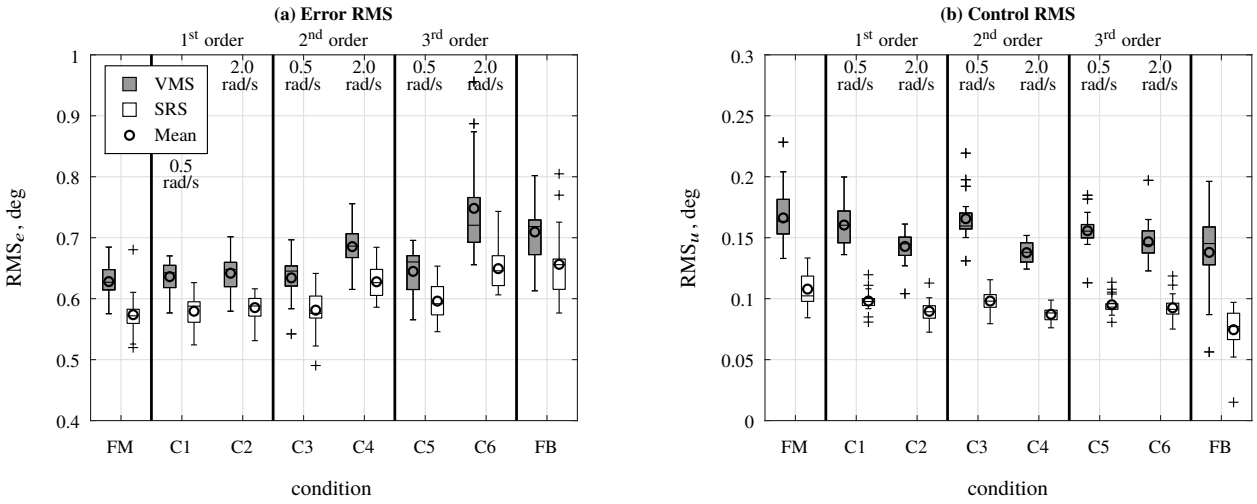

Fig. 5 Pilot tracking performance and control activity.

Table 4 Mixed-effects model results for pilot tracking performance and control activity based on conditions C1 to C6.

Measure	Simulator		O_{mf}		ω_{mf}		$O_{mf} \times \omega_{mf}$		Simulator × O_{mf}		Simulator × ω_{mf}		Simulator × $O_{mf} \times \omega_{mf}$	
	$\chi^2(1)$	p	$\chi^2(2)$	p	$\chi^2(1)$	p	$\chi^2(2)$	p	$\chi^2(2)$	p	$\chi^2(1)$	p	$\chi^2(2)$	p
RMS$_e$	6.28	0.01	33.32	<0.01	57.08	<0.01	30.03	<0.01	1.00	0.61	0.29	0.59	1.85	0.40
RMS$_u$	6.19	0.01	0.08	0.96	52.41	<0.01	11.32	<0.01	0.13	0.94	2.27	0.13	2.08	0.35

= significant ($p < 0.05$) = not significant ($p \geq 0.05$)

A. Performance and Control Activity

Fig. 5(a) shows the measured pilot tracking performance in terms of RMS_e. The best tracking performance – with an average RMS_e of 0.63 and 0.57 deg for the VMS and SRS data, respectively – was observed in the conditions with full motion (FM) and with a first-order motion filter and a low break frequency (C1). The worst performance (0.72 and 0.65 deg on average for VMS and SRS, respectively) was observed in the conditions without motion (FB) and with a third-order motion filter and a high break frequency (C6). Both simulators showed similar differences in tracking performance caused by different motion filter orders and break frequencies (C1-C6), which is supported by the insignificant interaction terms of O_{mf} and ω_{mf} with "Simulator", see Table 4. The mixed-model analysis showed a significant interaction between O_{mf} and ω_{mf}. Post-hoc analysis with pairwise comparisons indicated that for ω_{mf} = 2.0 rad/s performance significantly degraded (increase in RMS_e) with increasing filter order with significant differences between C2 and C4 ($p < 0.01$) and C4 and C6 ($p < 0.01$). For ω_{mf} = 0.5 rad/s, no significant differences between filter orders were found (C1 and C3, p = 1.00, and C3 and C5, p = 0.58). For O_{mf} = 1, no significant differences between the two break frequencies were found (C1 vs. C2, p = 0.99), while for O_{mf} = 2 and O_{mf} = 3 performance was significantly lower for ω_{mf} = 2.0 rad/s compared to ω_{mf} = 0.5 rad/s (C3 vs. C4, $p < 0.01$, and C5 vs. C6 $p < 0.01$). The mixed-model analysis also detected significant differences in tracking performance between the two simulators with pilots in the SRS performing consistently better (10% lower RMS_e on average).

Fig. 5(b) shows the pilot control activity measure, RMS_u. In the no-motion condition (FB) RMS_u was lowest on average for both simulators at 0.14 and 0.075 deg for the VMS and SRS, respectively . In the full-motion condition FM, the average RMS_u was highest (0.17 and 0.11 for VMS and SRS, respectively), indicating that pilots controlled more actively with full motion. The interactions of simulator with O_{mf} and ω_{mf} were not significant (Table 4), indicating similar effects of motion filter order and break frequency in both the VMS and SRS. The mixed-effects model analysis showed a significant interaction between filter order O_{mf} and filter break frequency ω_{mf}, see Table 4. Pairwise comparisons found no significant differences between filter orders for both ω_{mf} = 0.5 rad/s and ω_{mf} = 2.0 rad/s (C1 vs. C3, p = 0.84, C3 vs. C5, p = 0.21, C2 vs. C4, p = 0.55, C4 vs. C6, p = 0.11). Control activity was significantly lower under ω_{mf} = 2.0 rad/s compared to ω_{mf} = 0.5 rad/s for O_{mf} = 1 and O_{mf} = 2 (both $p < 0.01$); however, no significant differences between break frequencies were found for O_{mf} = 3 (p = 0.16). Finally, as is evident from Fig. 5(b) and Table 4, significant differences in RMS_u were found between the two simulators. In the VMS, control activity was found to be 64% higher on average, indicative of a more high-gain control strategy. This difference is attributed to either differences in the pilot population in both simulators, or the 3.9 cm longer arm of the sidestick used in the VMS (see Appendix A). As we considered the angular displacement of the stick as our control input u, a longer arm implied slightly larger movement at the grip point was required for the same input to $H_\theta(s)$ (i.e., additional 3.4 mm for a 5 deg input), which could have induced a more high-gain control strategy.

B. Pilot Model Equalization Parameters

As explained in Section III.E, the VAF was calculated to assess the quality of fit of the estimated pilot models. A VAF of 100% signifies that the corresponding pilot model was able to perfectly explain all the variance in the measured pilot control signal u. For the collected experiment data, the VAF values were 84.0% and 85.3% (ranging between 70% and 92%) on average for the VMS and SRS data, respectively, which indicates accurate pilot model fits, in line with results of previous experiments [35, 38].

Fig. 6 shows the estimated pilot equalization parameters (K_v, K_m, and T_L), as well as the corresponding control signal contribution fraction $\sigma^2_{u_m}/\sigma^2_{u_v}$. Fig. 6(a) shows the pilot model visual gain K_v. A higher K_v indicates that pilots responded with larger inputs to visual cues. Consistent with the RMS_u results in Fig. 5(b), 30% lower visual gains were found in the SRS compared to the VMS. This difference between simulators was, however, not found to be statistically significant, see Table 5. Fig. 6(a) and Table 5 show a significant interaction effect between O_{mf} and ω_{mf}. While with the low break frequency K_v remains approximately constant over the different filter orders (C1 vs. C3, $p = 1.0$, and C3 vs. C5, $p = 1.0$), for $\omega_{mf} = 2.0$ rad/s average decreases in K_v of 30% and 14% were found for the VMS and SRS, respectively (C2 vs. C4, $p = 0.05$ and C4 vs. C6, $p < 0.01$). No significant differences were found between the two break frequencies for the first-order filter ($p = 0.43$), but significant differences between the break frequencies were found for the second- and third-order filters ($p < 0.01$). The differences between break frequencies for the third-order filter were bigger for the VMS pilots, as indicated by the significant interaction of "Simulator" with O_{mf} and with the interaction of O_{mf} and ω_{mf} (Table 5).

Fig. 6(b) shows the estimated pilot model motion gains K_m. A higher K_m indicates that pilots respond with larger inputs to motion cues. Consistent with the results for K_v, the motion gains for the VMS pilots were higher (45% on average) than for the SRS pilots, but this difference was not statistically significant, see Table 5. In addition, the motion gain K_m showed similar relative trends for the different motion filter orders and break-frequencies in both simulators, i.e., the interactions of "Simulator" with O_{mf} and ω_{mf} were not significant, see Table 5. Just as for K_v, increased O_{mf} and ω_{mf} also resulted in significant reductions in K_m – 13% and 22% on average, respectively – meaning that the pilots responded less strongly to motion information in conditions with high O_{mf} or high ω_{mf}. Also, a significant interaction between O_{mf} and ω_{mf} was observed. For $\omega_{mf} = 0.5$ rad/s, no significant differences were found between the filter orders. For $\omega_{mf} = 2.0$ rad/s, a significant differences were found between $O_{mf} = 1$ and $O_{mf} = 2$ ($p < 0.01$); however, no significant differences were found between $O_{mf} = 2$ and $O_{mf} = 3$ ($p = 0.88$). In addition, no significant differences were found between the two break frequencies for $O_{mf} = 1$ ($p = 0.16$). For $O_{mf} = 2$ and $O_{mf} = 3$, the pilot motion gain was significantly lower for $\omega_{mf} = 2.0$ rad/s compared to $\omega_{mf} = 0.5$ rad/s ($p < 0.01$).

Fig. 6(c) shows the estimated pilot model visual lead time constants T_L. A higher lead time constant indicates that pilots use more visual lead to control the aircraft. As expected [41], in the no-motion FB condition the medians of the data of both simulators (VMS: 1.47 s, SRS: 1.25 s) were found to consistently approximate the inverse of the short

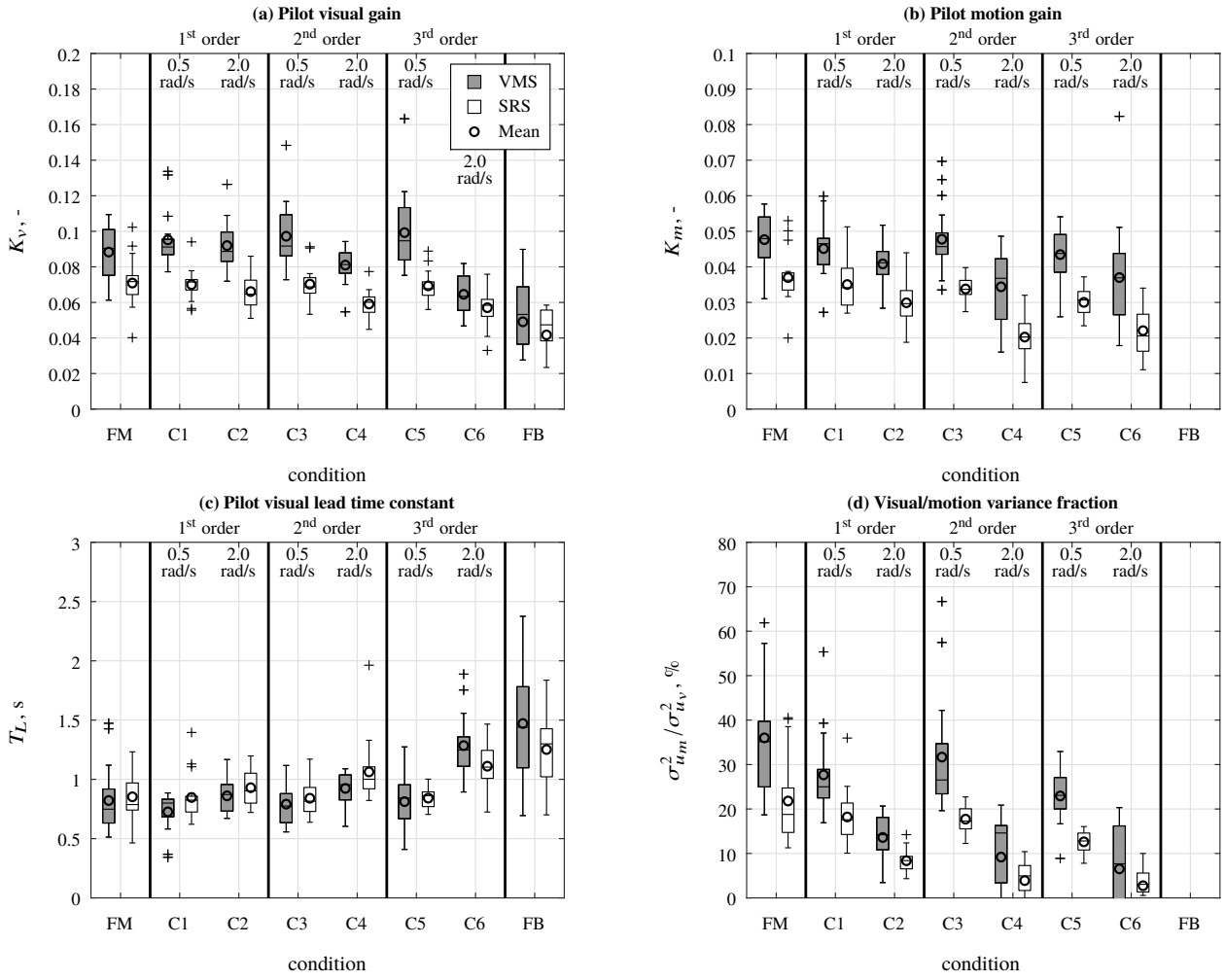

Fig. 6 Pilot equalization parameters and visual/motion variance fraction.

Table 5 Mixed-effects model results for pilot equalization parameters based on conditions C1 to C6.

Measure	Simulator		O_{mf}		ω_{mf}		$O_{mf} \times \omega_{mf}$		Simulator × O_{mf}		Simulator × ω_{mf}		Simulator × $O_{mf} \times \omega_{mf}$	
	$\chi^2(1)$	p	$\chi^2(2)$	p	$\chi^2(1)$	p	$\chi^2(2)$	p	$\chi^2(2)$	p	$\chi^2(1)$	p	$\chi^2(2)$	p
K_v	2.63	0.10	14.15	<0.01	55.22	<0.01	18.85	<0.01	6.80	0.03	0.71	0.40	7.42	0.02
K_m	3.64	0.06	9.02	0.01	45.63	<0.01	9.97	0.01	2.20	0.33	0.13	0.72	0.06	0.97
T_L	0.84	0.36	16.89	<0.01	67.73	<0.01	16.14	<0.01	7.51	0.02	4.72	0.03	3.77	0.15
$\sigma^2_{u_m}/\sigma^2_{u_v}$	4.03	0.04	17.40	<0.01	211.89	<0.01	13.40	<0.01	2.17	0.34	1.36	0.24	1.02	0.60

= significant ($p < 0.05$) = not significant ($p \geq 0.05$)

period frequency of the controlled dynamics: $1/\omega_{sp}$ = 1.4509 s. Fig. 6(c) and Table 5 indicate a significant interaction effect between O_{mf} and ω_{mf}. While with the low break frequency T_L remains approximately constant over the different filter orders (C1 vs. C3, p = 0.99, and C3 vs. C5, p = 1.0) and for ω_{mf} = 2.0 rad/s no differences between the first- and second-order filters were found, significant increases in T_L of 49% and 19% between the second- and third-order filters were found for the VMS and SRS, respectively (C2 vs. C4, p = 0.14, and C4 vs. C6, p < 0.01). Significant differences between the break frequencies were found for all filter orders (C1 vs. C2, p = 0.04, and C3 vs. C4 and C5 vs. C6, p < 0.01). Overall, no significant differences in T_L values between both simulators were found, see "Simulator" effect in Table 5. While the data of both simulators show similar trends across conditions, still a significant interaction of "Simulator" with O_{mf} and with ω_{mf} was found. Fig. 6(c) illustrates consistently stronger increases in T_L for the VMS data due to higher O_{mf} and ω_{mf}, which shows that the adaptation of T_L in both simulators was slightly different.

Fig. 6(d) shows the ratio of the variance of the part of the pilot model output originating from the motion perception channel ($\sigma^2_{u_m}$) and the variance of the model output coming from the visual perception channel ($\sigma^2_{u_v}$), see Fig. 1, as a measure of how much motion was used by pilots. A higher variance fraction indicates that more motion was used. A variance fraction of 100% indicates that the variances of the motion and visual contributions to u were equal. The values ranged from 28.9% on average for the FM condition to 4.6% on average for C6. For the no-motion condition (FB) no data were available, as the motion channel of the pilot model was not estimated. A significant interaction between O_{mf} and ω_{mf} was present. There were no significant differences between a first-order and a second-order filter for ω_{mf} = 0.5 rad/s (p = 0.99), but there were a significant differences between a second- and third-order filter (p < 0.01). For ω_{mf} = 2.0 rad/s, significant differences were found between a first- and second-order filter (p < 0.01), and a second- and third-order filter (p < 0.01). $\sigma^2_{u_m}/\sigma^2_{u_v}$ values were significantly lower for ω_{mf} = 2.0 rad/s compared to ω_{mf} = 0.5 rad/s for all filter orders (p < 0.01). Overall, Fig. 6(d) also shows significant differences between the VMS and SRS, as confirmed by the significant main effect of "Simulator" in Table 5, with the VMS pilots relying more on the provided motion feedback. Both simulators did show similar trends in $\sigma^2_{u_m}/\sigma^2_{u_v}$, however, as indicated by the lack of significant interactions of "Simulator" with O_{mf} and ω_{mf}, see Table 5.

C. Pilot Model Limitation Parameters

Fig. 7 shows the estimated pilot perception delays and neuromuscular parameters, with Table 6 listing the corresponding statistical analysis results. Fig. 7(a) depicts the visual perception time delays τ_v. Pilot's visual time delay remained approximately constant over the different conditions in each simulator. The mixed-effects models indicated no significant interactions of "Simulator" with O_{mf} and ω_{mf}, see Table 6. Furthermore, no significant interaction between O_{mf} and ω_{mf} was found. However, there was a small significant main effect of motion filter break frequency (Table 6). The visual time delay was found to be marginally higher for the conditions with ω_{mf} = 2.0 rad/s. A significant difference between simulators was present, see Table 6, with visual time delays being consistently around

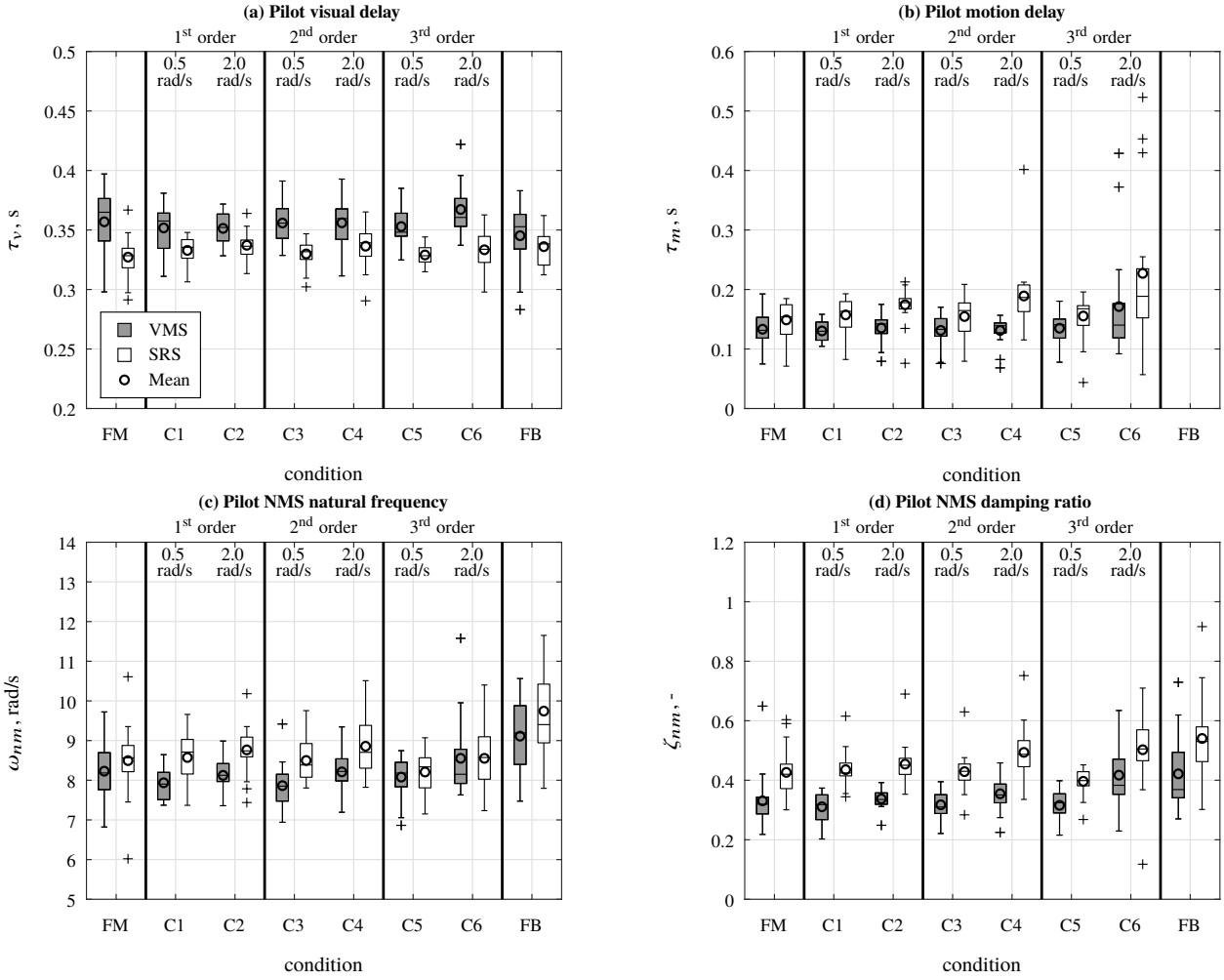

Fig. 7 Pilot delays and neuromuscular model parameters.

Table 6 Mixed-effects model results for pilot delays and neuromuscular parameters based on conditions C1 to C6.

Measure	Simulator		O_{mf}		ω_{mf}		$O_{mf} \times \omega_{mf}$		Simulator × O_{mf}		Simulator × ω_{mf}		Simulator × $O_{mf} \times \omega_{mf}$	
	$\chi^2(1)$	p	$\chi^2(2)$	p	$\chi^2(1)$	p	$\chi^2(2)$	p	$\chi^2(2)$	p	$\chi^2(1)$	p	$\chi^2(2)$	p
τ_v	8.54	<0.01	0.52	0.77	4.63	0.03	1.78	0.41	5.04	0.08	0.01	0.94	2.77	0.25
τ_m	0.62	0.43	1.16	0.56	3.52	0.06	0.43	0.81	0.68	0.71	0.42	0.52	0.40	0.82
ω_{nm}	1.37	0.24	0.05	0.97	11.76	<0.01	1.03	0.60	9.22	0.01	0.04	0.83	0.13	0.94
ζ_{nm}	5.93	0.01	4.46	0.11	41.13	<0.01	16.78	<0.01	2.07	0.36	0.00	0.98	0.34	0.84

= significant ($p < 0.05$) = not significant ($p \geq 0.05$)

22 ms higher in the VMS.

The estimated pilot motion perception time delays τ_m are depicted in Fig. 7(b). Like the visual time delay τ_v, the motion time delay τ_m remained approximately constant over the different motion conditions. No significant interactions of simulator with O_{mf} and ω_{mf} were found and no significant interactions between O_{mf} and ω_{mf}. There were also no statistically significant differences in pilot motion time delay between simulators, see Table 6. Compared to the values of $\tau_m \approx 0.15$-0.2 s reported previous studies [18, 35, 40], some outliers were found for C4 and C6. These outliers indicate that due to the less consistent use of motion feedback (i.e., low $\sigma^2_{u_m}/\sigma^2_{u_v}$) in these low-fidelity motion conditions, τ_m could not be estimated accurately. Similar issues with outlying pilot delay estimates have been reported for earlier experiments [7, 46, 47].

The neuromuscular frequency ω_{nm} is shown in Fig. 7(c). A lower ω_{nm} indicates more relaxed arm dynamics (less muscle tension) and a reduced bandwidth of the pilot's neuromuscular frequency response. While ω_{nm} values were of equivalent magnitude in both simulators, a significant interaction was found between "Simulator" and O_{mf}, indicating different effects of motion filter order in the two simulators, see Table 6. In the VMS, a small increase in ω_{nm} was found with increasing filter order, while in the SRS a small decrease was measured. Overall, ω_{nm} was also found to be significantly affected by ω_{mf}, as the neuromuscular frequency was consistently higher with higher break frequencies. No significant interaction of the effects of O_{mf} and ω_{mf} was found.

Fig. 7(d) shows the estimated neuromuscular damping ratio ζ_{nm}. A lower ζ_{nm} indicates that the pilot's neuromuscular dynamics are less well-damped. Fig. 7(d) shows that the neuromuscular damping was consistently around 0.1 higher in the SRS than in the VMS, a significant effect, see Table 6. As this difference is consistent across all conditions, including the no-motion condition FB, the difference in motion systems was not the source of this difference across simulators. This result is consistent with the longer arm of the sidestick in the VMS (see Appendix A), resulting in a less well-damped neuromuscular response. In both simulators similar trends were present, as indicated by the lack of statistically significant interaction terms for "Simulator" in Table 6. The mixed-model analysis found a significant interaction between motion filter order and break frequency, see Table 6. No direct effect of O_{mf} on ζ_{nm} was found for $\omega_{mf} = 0.5$ rad/s (C1 vs. C3, $p = 1.00$, and C3 vs. C5, $p = 0.87$) and $\omega_{mf} = 2.0$ rad/s (C2 vs. C4, $p = 0.44$, and C4 vs. C6, $p = 0.05$). While no significant differences between the two break frequencies were found for $O_{mf} = 1$ ($p = 0.62$), ζ_{nm} was significantly higher for $\omega_{mf} = 2.0$ rad/s using second and third order filters ($p < 0.01$).

D. Open-loop Dynamics

Fig. 8 shows the target and disturbance open-loop crossover frequencies and phase margins for both the VMS and SRS experiments. The corresponding statistical analysis results are listed in Table 7. The open-loop target crossover frequency $\omega_{c,t}$ is shown in Fig. 8(a). It was found to increase up to 0.2 rad/s with increasing filter order and with increasing break frequency, both statistically significant effects, see Table 7. As the increase in $\omega_{c,t}$ due to ω_{mf} was

consistent for all O_{mf} settings, no significant interaction between the effects of order and break frequency was found. While $\omega_{c,t}$ was found to be on average 0.2 rad/s higher for the VMS pilots, a result that is consistent with the pilot model gains presented in Fig. 6, this between-simulator difference was not statistically significant. Furthermore, the variation across conditions was found to be equivalent in both simulators, which is confirmed by the lack of significant interaction effects with "Simulator" in our mixed-effects model analysis, see Table 7.

The estimated open-loop disturbance crossover frequencies $\omega_{c,d}$ are shown in Fig. 8(b). A significant interaction effect was found between O_{mf} and ω_{mf}, see Table 7. For ω_{mf} = 0.5 rad/s, no significant differences were found between the first- and second-order filters ($p = 0.95$), but $\omega_{c,d}$ was significantly lower with a third-order filter compared to a second-order filter ($p = 0.02$). For ω_{mf} = 2.0 rad/s, $\omega_{c,d}$ was significantly lower for $O_{mf} = 2$ compared to $O_{mf} = 1$ ($p < 0.01$), but no significant differences were found between $O_{mf} = 2$ and $O_{mf} = 3$ ($p = 0.99$). $\omega_{c,d}$ was significantly lower for the higher break frequency for all filter orders (p <0.01). Furthermore, the main effect of "Simulator" was not significant, indicating that there were no absolute differences between the simulators, see Table 7. The mixed-model analysis detected no significant interaction effects of "Simulator" with O_{mf} and ω_{mf}, indicating that the relative trends introduced by O_{mf} and ω_{mf} were the same in both simulators (Table 7).

Fig. 8(c) shows the open-loop target phase margins $\varphi_{m,t}$. The main effect of "Simulator" was statistically significant, as $\varphi_{m,t}$ was consistently lower, 4 deg on average, for the VMS data. Despite this offset, both simulators showed similar trends in the data, i.e., no significant interactions of "Simulator" with O_{mf} and ω_{mf} were found, see Table 7. A significant interaction was found between O_{mf} and ω_{mf} (Table 7). Post-hoc tests found similar differences between conditions compared to $\omega_{c,d}$. For ω_{mf} = 0.5 rad/s, no significant differences were found between $O_{mf} = 1$ and $O_{mf} = 2$ ($p = 1.00$), but $\varphi_{m,t}$ was significantly lower with $O_{mf} = 3$ compared to $O_{mf} = 2$ ($p < 0.01$). For ω_{mf} = 2.0 rad/s, $\varphi_{m,t}$ was significantly lower for $O_{mf} = 2$ compared to $O_{mf} = 1$ ($p = 0.02$), but no significant differences were found between $O_{mf} = 2$ and $O_{mf} = 3$ ($p = 0.88$). The target phase margins were significantly lower for the higher break frequency for all filter orders (p <0.01). Furthermore, for $\varphi_{m,t}$ a significant interaction between O_{mf} and ω_{mf} was found, as the reduction in phase margin varies marginally across the different O_{mf} settings.

Fig. 8(d) shows the open-loop disturbance phase margins $\varphi_{m,d}$. A significant interaction effect between O_{mf} and ω_{mf} was found. While with the low break frequency $\varphi_{m,d}$ remains approximately constant over the different filter orders (C1 vs. C3, $p = 1.0$ and C3 vs. C5, $p = 0.83$), for $\omega_{mf} = 2.0$ rad/s average decreases in $\varphi_{m,d}$ were found between the first- and the second-order filters ($p < 0.01$), but not between the second- and third-order filters ($p = 0.11$). No significant differences were found between the two break frequencies for the first-order filter ($p = 0.99$), but disturbance phase margins were significantly lower with the higher break frequency for the second- and third-order filters ($p < 0.01$). Consistent with the $\varphi_{m,t}$ data in Fig. 8(c), the disturbance phase margins are found to be 5 deg lower on average for all experiment conditions in the VMS data compared to the SRS data, a significant effect as can be verified from Table 7. As the effect of ω_{mf} on $\varphi_{m,d}$ was found to be slightly stronger in the SRS compared to the

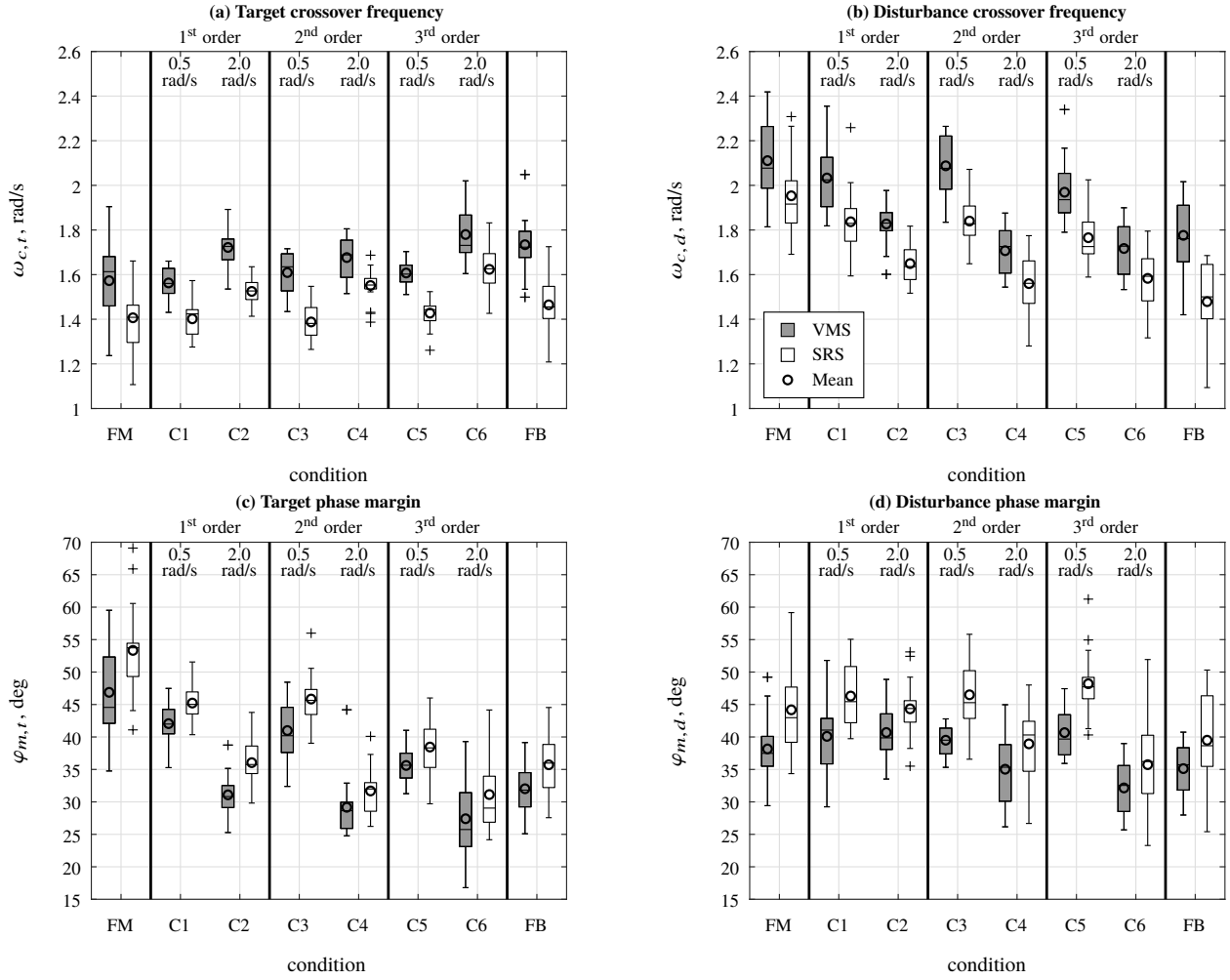

Fig. 8 Open-loop crossover frequencies and phase margins.

Table 7 Mixed-effects model results for crossover frequencies and phase margins based on conditions C1 to C6.

Measure	Simulator		O_{mf}		ω_{mf}		$O_{mf} \times \omega_{mf}$		Simulator × O_{mf}		Simulator × ω_{mf}		Simulator × $O_{mf} \times \omega_{mf}$	
	$\chi^2(1)$	p	$\chi^2(2)$	p	$\chi^2(1)$	p	$\chi^2(2)$	p	$\chi^2(2)$	p	$\chi^2(1)$	p	$\chi^2(2)$	p
$\omega_{c,t}$	2.75	0.10	10.12	0.01	111.51	<0.01	5.61	0.06	0.15	0.93	1.42	0.23	5.46	0.07
$\omega_{c,d}$	1.85	0.17	5.87	0.05	126.15	<0.01	10.42	0.01	0.44	0.80	3.19	0.07	0.92	0.63
$\varphi_{m,t}$	4.26	0.04	20.51	<0.01	168.44	<0.01	12.97	<0.01	0.29	0.86	0.01	0.92	2.24	0.33
$\varphi_{m,d}$	5.92	0.01	12.50	<0.01	51.47	<0.01	30.96	<0.01	0.16	0.92	5.42	0.02	0.17	0.92

= significant ($p < 0.05$) = not significant ($p \geq 0.05$)

VMS, a significant interaction between "Simulator" and ω_{mf} was found, see Table 7. However, despite this single significant "Simulator" interaction from the mixed-effect model analysis, overall it can be concluded that the variation in the open-loop parameters shown in Fig. 8 is highly consistent across simulators.

V. Discussion

For the experiment described in this paper, the first hypothesis H1 was that with increasing filter order O_{mf}, pilots would display worse performance and reduced use of motion feedback in their control strategy. The collected task performance data, with an increase in RMS_e for higher O_{mf} confirm this performance degradation. A significant increase in visual lead time T_L further suggests that pilots relied more on visual cues to control the aircraft. Both effects were stronger for the higher motion filter break frequency setting $\omega_{mf} = 2.0$ rad/s. A significant decrease of 13% in pilot motion feedback gain K_m with increasing O_{mf} was found, which confirms that pilots responded less strongly to the motion signal. The ratio of control signal variances $\sigma^2_{u_m}/\sigma^2_{u_v}$ also decreased significantly for higher motion filter orders. Finally, in both simulators the crossover frequencies $\omega_{c,t}$ and $\omega_{c,d}$ showed the expected increase and decrease, respectively, when increasing O_{mf}. The corresponding phase margins, decreased significantly with up to 5 deg on average, indicating that pilots controlled with reduced stability margins with higher motion filter orders.

This result is in line with available motion fidelity criteria [32, 44], which to our knowledge were not formally validated for a wide range of filter orders as tested in the current experiment, and previous experiments where motion conditions of varying fidelity were tested, e.g., [17]. Furthermore, as expected, the experimental conditions with $\omega_{mf} = 2.0$ rad/s showed stronger effects of increasing O_{mf}, which was also evident from the significant interactions between O_{mf} and ω_{mf} observed in many of the considered dependent measures. With every increasing order, the motion filter adds up to 90 degrees of additional phase lead to the supplied motion feedback. In the conditions with $\omega_{mf} = 2.0$ rad/s this phase distortion was present over a wider frequency range, i.e., it overlapped more with the bandwidth of the task performed, as compared to $\omega_{mf} = 0.5$ rad/s. This explains why increasing O_{mf} would have impacted the pilots more. Overall, the obtained results show that a first-order motion filter resulted in behavior (i.e., K_v, K_m, and T_L, but also the target and disturbance crossover frequencies and phase margins) that most closely matched the full-motion condition. However, when ω_{mf} was sufficiently low, the filter order did not influence pilot behavior, which also resulted in better pilot performance. Overall, based on these results of O_{mf} and its interaction with ω_{mf} observed in our data, Hypothesis H1 was accepted.

It should be noted that due to our choice to focus on measuring the interaction effects between O_{mf} and ω_{mf} in a factorial experiment design, the third-order filters tested in our experiment were of lower fidelity than those typically applied in flight simulators. In practice, the additional first-order filter break frequency in a third-order motion filter is generally set to a much lower value than the second-order ω_{mf} (i.e., 0.2 or 0.3 rad/s) to provide additional washout of only the slowest aircraft accelerations [13, 25]. Consequently, the pilot control behavior changes that would be

expected for such practical third-order motion filters are expected to be closer to the $O_{mf} = 2$ results presented in this paper, than those of our $O_{mf} = 3$ conditions.

The second hypothesis H2 was that with increasing motion filter break frequency ω_{mf}, pilots would show similar effects as with an increasing motion filter order O_{mf}, but that between the considered settings of 0.5 and 2 rad/s the effects of ω_{mf} would be stronger. Except for the motion perception delay τ_m, all dependent measures were indeed significantly affected by the increase in ω_{mf}. In the data of both simulators, pilots showed a higher RMS_e (VMS: 8%, SRS: 6%) with increasing ω_{mf}. Furthermore, the variance fraction $\sigma^2_{u_m}/\sigma^2_{u_v}$ indicated that with increasing ω_{mf}, less use was made of motion feedback. The motion gain K_m supported this with a similar significant decrease. The visual equalization parameters K_v and T_L showed a significant decrease and an increase for higher ω_{mf}, respectively. This again indicates that pilots used more visual cues to generate lead, while controlling with smaller gains. All four open-loop variables ($\omega_{c,t}$, $\omega_{c,d}$, $\varphi_{m,t}$, and $\varphi_{m,d}$) showed significant effects of ω_{mf}, consistent with a lower dependence on the motion channel of the pilot model. Thus, Hypothesis H2 was accepted, as pilots showed the expected decrease in use of the motion channel and a reduced performance with increasing ω_{mf}. These findings are consistent with previous experiments [16, 17] and suggest that the Sinacori-Schroeder motion fidelity criterion [32, 44] is a reasonable predictor for the combined effects of ω_{mf} and O_{mf}.

The experiment described in this paper measured the effects of variations in motion cueing quality on pilot behavior in tracking tasks with only a heads-down visual display and physical motion feedback. Compared to fully realistic simulator environments, our setup thus did not include an immersive out-of-the-window visual and peripheral visual cues, which are known to have a similar, yet less pronounced, effects on pilot control behavior as physical motion feedback [48–51]. With an out-of-the-window view, pilots can also perceive "lead" information on the vehicle motion from the available peripheral visual cues, especially since the perceptual delay on this peripheral visual lead is lower than that for lead generated from central visual cues (i.e., from $H_{p_{vis}}$) [48]. While this means that measured effects of degraded simulator motion fidelity on pilot control behavior are likely less strong in the presence of peripheral visual cues, key motion feedback effects such as reducing control gains and increased visual equalization have been found to persist with out-of-the-window visual cues [50]. While for pitch control the effects of peripheral visual cues are likely less strong than earlier experiments have reported for roll [48, 49, 51] and yaw [50], verification of the potential effects of out-of-the-window visual feedback for the data presented in this paper would require further experiments, where peripheral visual feedback is explicitly included as an additional factor.

Even though meticulous attention was paid to match relevant hardware systems (and equivalent dynamics) between the VMS and SRS for this experiment, half of our dependent measures showed clear offsets between both simulators, resulting in some significant mixed-effect model results for our between-group "Simulator" variable. Therefore, Hypothesis H3, which stated that no between-simulator offsets would be present in the experiment data, is rejected. These differences between simulators are partially explained by the substantially more high-gain control strategy used

by some of the VMS pilots, which had a direct effect on RMS_u, K_v, K_m, $\sigma^2_{u_m}/\sigma^2_{u_v}$, and the open-loop parameters. Part of these differences are likely also attributable to the difference in sidestick arm length between both simulators. The sidestick used in the VMS had a 3.9 cm longer arm, which thus required slightly larger inputs at the stick grip for the same angular stick input u and might have induced some pilots to control with higher control intensity (RMS_u). Furthermore, the armrest in the VMS was covered with a canvas fabric, which allowed pilots' arm to slide freely over the armrest. The armrest in the SRS, however, was covered in artificial leather, which impeded free movement of the arm to a certain degree. The difference in ζ_{nm} in the no-motion condition FB supported both these claims. Finally, perhaps the most likely explanation for the between-simulator offsets found in our experiment is the use of different participants in both simulators. In the end, more of the VMS pilots had prior experience in tracking tasks, whereas the SRS pilots generally performed such a task for the first time. Even though training runs ensured that all pilots showed equally constant individual performance (as assessed with RMS_e) and the less-experienced SRS pilots in fact had superior task performance, any differences between the VMS and SRS participants also would have contributed to the observed "Simulator" effects. Unfortunately, as it will be impossible to eliminate *all* hardware differences and ensure a fully identical participant pool in comparison studies, no between-simulator comparison experiment based on measurements of pilot control behavior will be completely free of such direct between-simulator differences.

Despite the fact that half of our pilot control behavior related metrics showed offsets between the data collected in both simulators, with only a few minor exceptions, all trends in these metrics between the tested motion conditions (i.e., the effects of the within-subjects variables O_{mf} and ω_{mf}) were very much consistent across simulators; Hypothesis H4 is accepted. The only parameters for which significant interactions of simulator with O_{mf} and ω_{mf} were found, were K_v, T_L, ω_{nm}, and $\varphi_{m,d}$. Although these interactions were significant, the magnitude of the differences in the trends between simulators was, in fact, quite small. In addition, including the reference motion conditions FB and FM in the second mixed-effects model analysis revealed interactions of simulator with motion condition for RMS_u and T_L only (Appendix B). Overall, it can be concluded that the relative trends between motion conditions are accurately reproducible across matched simulators, even in the presence of clear between-simulator offsets in the results. This result suggests that future experiment replications of behavior measured in flight simulators, and efforts at behavioral simulator benchmarking, should perhaps not focus on absolute standards for behavioral parameters, but rather on whether relative differences between different simulator (cueing) settings are replicated accurately. For example, a simulator's capacity for inducing a sufficiently strong difference in key pilot behavior metrics (e.g., K_v and T_L) between performing a task in fixed-base and full-motion conditions is likely a more robust predictor of a simulator's motion fidelity than the behavior shown with full motion itself. Future work should test this in an experiment that, unlike the experiment described in this paper, compares pilot behavioral metrics, obtained from the same pilots, between (deliberately) mismatched simulators.

VI. Conclusions

This paper used a pitch tracking experiment with 38 general aviation pilots to investigate the effects of different motion filter orders and break frequencies on pilot control behavior and performance. A full-factorial design with three motion filter orders and two motion filter break frequencies was used in addition to no-motion and full-motion reference conditions, resulting in eight motion conditions in total. The experiment was performed in the Vertical Motion Simulator (VMS) at NASA Ames Research Center and repeated in the SIMONA Research Simulator (SRS) at Delft University of Technology in order to verify the results across simulators. The motion systems of both simulators were equalized using shaping filters, the head-down displays used the same graphics, and the sidesticks used the same force-displacement settings. The contribution of motion feedback to pilots' control strategy was found to decrease for higher motion filter orders, leading to significantly lower tracking performance. As expected, the effects of increasing motion filter break frequency from 0.5 to 2 rad/s on pilot manual control behavior and performance were equivalent to those of increasing the motion filter order, albeit stronger than observed between the different tested filter orders. Overall, the results also confirm statistically significant interaction effects between filter order and break frequency settings in nearly all considered pilot behavior metrics. Matching objective predictions of this interaction by the Sinacori-Schroeder motion fidelity criterion, the effects of filter order variations on pilot control behavior were found to be strongly amplified at high filter break frequency settings. Even with the meticulous attention that was paid to equalizing the simulator systems and the pilot population in the combined VMS and SRS experiment described in this paper, the pilot behavior measures for the same motion conditions between both simulators show (considerable) differences in absolute value. However, the relative trends replicated well across the two simulators. These results indicate that pilot-in-the-loop experiment replications should focus on reproducing relative effects between multiple tested conditions, rather than aiming to achieve an exact match in pilot control behavior and performance as an absolute indicator of, for example, simulator fidelity.

Acknowledgments

The authors thank the simulation engineers and motion operators at NASA Ames SimLabs who contributed to the experiment in the VMS. We especially thank Steve Norris for his valuable contributions. We also thank the pilots who participated in the SRS and VMS experiments. Finally, we thank Steve Beard for making the experiment in the VMS possible.

A. Simulator Matching

Fig. 1 shows a schematic overview of the control task performed in both the VMS and the SRS simulators for the experiment described in this paper. In both simulators, pilots received cues from several simulator hardware systems (depicted in light gray boxes in Fig. 1): the motion system, visual system, and control device (inceptor). The following

sections describe how these systems and their cues were matched across both simulators.

A. Cockpit Dimensions

Figs. 9(a) and (b) show the locations of the control device and the head-down display with respect to the eye reference point (ERP) in both simulators. The only relevant difference between simulators was a vertical offset between the ERP and screen positions. The SRS ERP was 5 cm higher relative to the bottom of the screen compared to the VMS. The position of the sidesticks with respect to the ERP differed 2 cm. These differences were considered marginal, and no corrections were made.

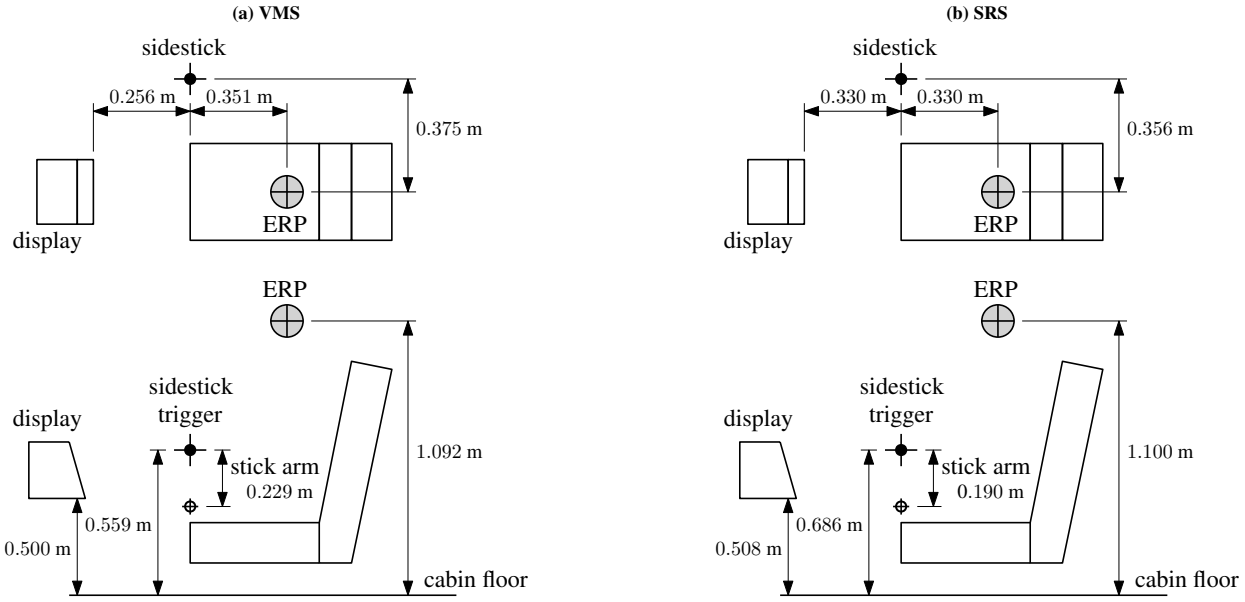

Fig. 9 Simulator cockpit dimensions.

B. Motion System

The VMS has an uncoupled, six-degree-of-freedom motion system [52]. The vertical and large horizontal motion axes are driven by electric motors, while the short horizontal and rotational axes are driven by hydraulic actuators [52]. The VMS was configured to use the large horizontal motion axis for surge motion, instead of the normal configuration in which the large axis is used for sway motion. The SRS has a hexapod motion system with six linear hydraulic actuators that have an operational stroke length of 1.15 m [14]. Table 8 provides the operational motion ranges for each degree-of-freedom of both simulators. For the pitch and heave motion axes of the VMS the equivalent time delays are 47 and 67 ms [6, 53], respectively. For the hexapod motion system of the SRS, the equivalent time delay is 25 ms in all axes [54].

To match the pitch and heave motion system dynamics ($H_{mot}(s)$ in Fig. 1) between simulators, motion shaping

Table 8 Simulator motion system ranges for the VMS [52, 53] and SRS [54, 55].

Axis	VMS	SRS
Surge	9.14 m*	2.24 m
Sway	1.83 m*	2.06 m
Heave	13.41 m	1.31 m
Roll	27.50 deg	51.80 deg
Pitch	27.50 deg	48.00 deg
Yaw	38.96 deg	83.20 deg

*Note that the VMS surge/sway workspace represents the motion system configuration for the current experiment, not the normal VMS workspace.

filter dynamics were estimated. These shaping filters were placed between the aircraft dynamics output and the motion filters on the SRS, see Fig. 1. The shaping filters had the following form [31]:

$$H_{shp}(s) = H_{SRS}^{-1}(s) \cdot H_{VMS}(s), \tag{9}$$

where $H_{SRS}(s)$ and $H_{VMS}(s)$ represent the unaltered motion frequency responses of the SRS and VMS, respectively. Objective Motion Cueing Test (OMCT) [8] measurements were performed for the full-motion condition FM (see Table 2) on both simulators prior to the experiment to determine the unaltered motion responses at twelve discrete OMCT measurement frequencies. Then, the transfer function model of Eq. (10) was fitted through the twelve OMCT frequency response measurement points using a quadratic cost function [31]:

$$H_{SRS,VMS}(s) = \frac{A \cdot s^2}{B \cdot s^2 + C \cdot s + D} \cdot e^{-E \cdot s} \tag{10}$$

The resulting simulator motion dynamics for the pitch degree-of-freedom of the VMS and SRS were:

$$H_{VMS_q}(s) = \frac{0.893 \cdot s^2}{0.916 \cdot s^2 + 0.254 \cdot s + 0.035} \cdot e^{-0.045 \cdot s} \tag{11}$$

$$H_{SRS_q}(s) = \frac{0.908 \cdot s^2}{0.891 \cdot s^2 + 0.259 \cdot s + 0.035} \cdot e^{-0.026 \cdot s} \tag{12}$$

Equivalently, the simulator motion dynamics for the heave degree of freedom were:

$$H_{VMS_z}(s) = \frac{0.911 \cdot s^2}{0.883 \cdot s^2 + 0.280 \cdot s + 0.036} \cdot e^{-0.098 \cdot s} \tag{13}$$

$$H_{SRS_z}(s) = \frac{0.908 \cdot s^2}{0.900 \cdot s^2 + 0.256 \cdot s + 0.036} \cdot e^{-0.045 \cdot s} \tag{14}$$

The standard OMCT test signal amplitudes as specified in [8] did not fit in the available motion space as the motion

gain was unity in all motion conditions (K_{mf} = 1.0). Some of the lower frequencies exceeded workspace limits in the SRS and some of the higher frequencies exceeded velocity limits in the VMS. Hence, using simulations of the motion systems of both simulators, the amplitudes were modified according to Table 9. Two sets of OMCT tests were performed. First, the full-motion condition FM was used to estimate the shaping filters of Eqs. 11 to 14. Second, with the shaping filters present, the motion filter settings of motion condition C5 (see Table 2) were used for verification of the matching of the motion system dynamics.

Table 9 OMCT signal amplitudes.

	Standard OMCT amplitudes		FM OMCT amplitudes		C5 OMCT amplitudes	
Frequency	Linear	Rotational	Linear	Rotational	Linear	Rotational
[rad/s]	[m/s^2]	[deg/s^2]	[m/s^2]	[deg/s^2]	[m/s^2]	[deg/s^2]
0.100	1.000	0.060	0.010	0.060	0.500	1.000
0.158	1.000	0.150	0.010	0.150	0.500	1.000
0.251	1.000	0.251	0.020	0.251	0.250	0.251
0.398	1.000	0.398	0.050	0.398	0.250	0.398
0.631	1.000	0.631	0.050	0.631	0.250	0.631
1.000	1.000	1.000	0.100	1.000	0.500	1.000
1.585	1.000	1.585	0.500	1.585	1.000	1.585
2.512	1.000	2.512	1.000	2.512	1.000	2.512
3.981	1.000	3.981	1.000	3.500	1.000	3.500
6.310	1.000	6.310	1.000	6.000	1.000	6.000
10.000	1.000	10.000	1.000	7.000	1.000	7.000
15.849	1.000	10.000	1.000	7.000	1.000	7.000

Fig. 10 shows the unaltered motion responses of both simulators and the SRS response with the shaping filter included for FM. Fig. 11 shows the same for experimental condition C5. In both figures it can be seen that the shaping filter succeeds in matching the SRS to the original unaltered VMS motion response for the heave degree-of-freedom. The pitch degree-of-freedom showed equivalent results, which are not included here for brevity.

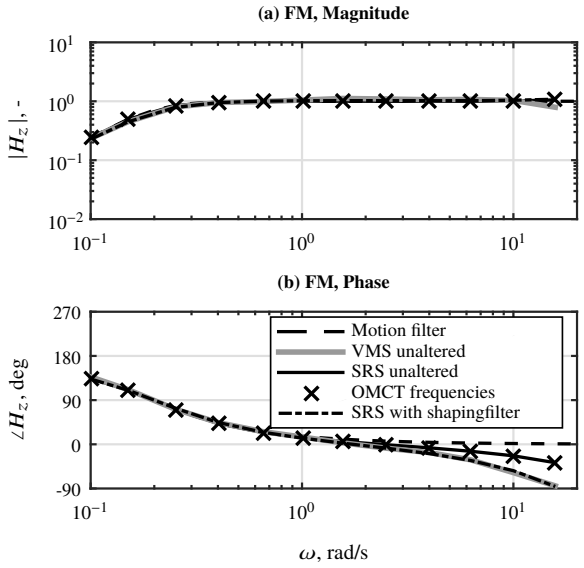

Fig. 10 Heave OMCT responses for FM.

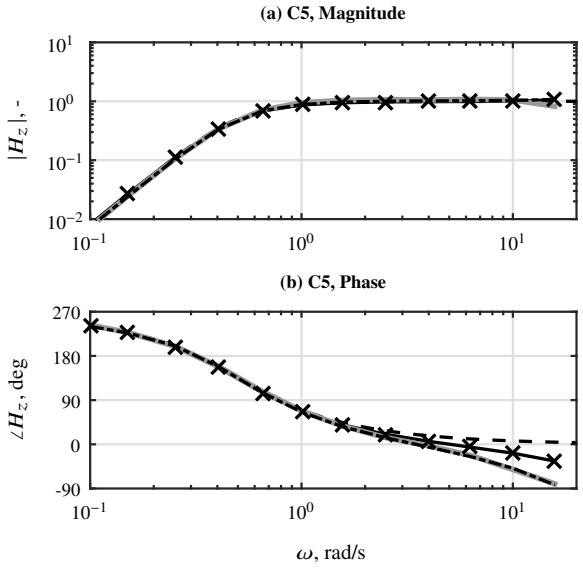

Fig. 11 Heave OMCT responses for C5.

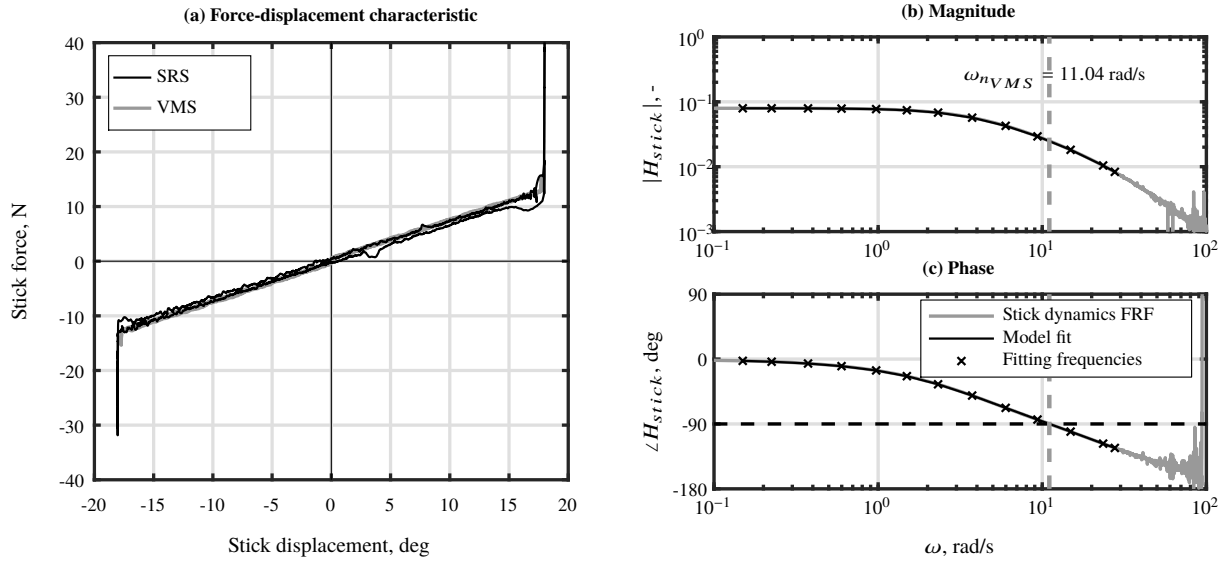

Fig. 12 Stick verification data.

C. Control Device

The VMS featured an electro-hydraulic McFadden control sidestick, whereas the SRS had an electrical Moog sidestick. Both simulators had an armrest installed. The armrest in the VMS was covered with a canvas fabric, which allowed the pilots' arms to slide relatively freely. The armrest in the SRS had an artificial leather covering, which prevented free movement of the arm to a certain degree. Table 10 presents the parameters of the sidestick used in the experiment. In both simulators the sidestick settings were set to these values, which were subsequently verified using a force-displacement plot in both simulators (Fig. 12(a)) and a frequency sweep in the SRS (Fig. 12(b) and (c)). The force-displacement plot in Fig. 12(a) allowed for verification of the gradient, the breakout, and the range of motion of both sidesticks. From the estimated frequency response data in Fig. 12(b) and (c), the natural frequency of the SRS stick was found by fitting a mass-spring-damper transfer function to the stick dynamics H_{stk} determined using the frequency sweep data. The stick dynamics were determined using the relation $H_{stk}(j\omega) = U(j\omega)/F(j\omega)$, where $U(j\omega)$ and $F(j\omega)$ are the Fourier-transformed control signal (i.e., stick position) and stick force, respectively. The natural frequency of the VMS sidestick was found to be $\omega_n = 11.04$ rad/s by manually adjusting a mechanical damping factor in the sidestick hardware and subsequently letting the stick oscillate in its natural frequency after a small perturbation. Fig. 12(b) and (c) show that the mass-spring-damper transfer function fitted on the SRS frequency sweep data crosses the -90 degrees phase line at $\omega_n = 11.08$ rad/s. One difference between the two sidesticks that could not be adjusted was the length of the stick arm, which was 0.229 m in the VMS and 0.190 m in the SRS, as measured from the rotation axis to the bottom of the trigger, Figs. 9(a) and (b). Furthermore, the design of the grip of both sticks was different.

Table 10 Overview of sidestick parameters.

Parameter	Unit	Value
Max Deflection	deg	± 18.0000
Force Gradient	N/deg	0.6987
Breakout Force	N	0.0000
Damping	Ns/deg	0.1747
Inertia	Ns^2/deg	0.0057
Natural Frequency	rad/s	11.0400

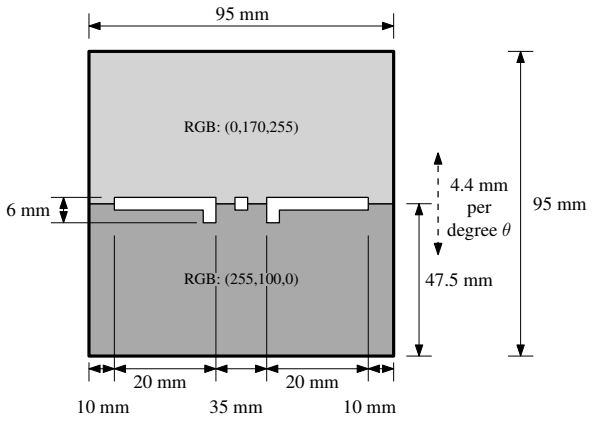

Fig. 13 Dimensions of the primary flight display.

D. Visual System

To eliminate effects of different out-of-the-window visual systems and simplify the replication of the experiment, only a head-down display was used for the experiment. The display graphics on the head-down displays in the VMS and SRS were generated from the same C/OpenGL code. The dimensions of the visual features were measured and adjusted such that the aircraft symbol and artificial horizon on the SRS display replicated those on the VMS display. Fig. 13 depicts the dimensions of the PFD on the screen.

In a previous simulator comparison, the dynamics of the displays were modelled as pure time delays [31]. The time delay of the visual display system in the VMS was measured using the Image Dynamic Measurement System (IDMS) [56] and was found to be 36.3 ms. The IDMS is based on detecting a change from black to white on the screen. It uses an instrument with a video input that measures the time it takes between the command being generated and the change to happen on the display. The total time delay of the visual display system in the SRS was measured using the Visual Delay Measurement System (VDMS) [57]. The VDMS test is based on a sinusoidal input signal on the pitch angle. The resulting image is sampled at twice the sinusoidal frequency through a set of fast LCD shutter glasses. The opening of the glasses can be shifted in time by an observer, who adjusts the offset until both zero-crossings are visible. The resulting time offset is used as an approximation of the display's latency. The procedure is repeated for three frequencies: 2, 4 and 8 Hz. Two different observers performed the procedure, resulting in an estimated visual delay of 33-39 ms. Because this fell within the same range of the VMS visual time delay, no adjustments were needed to match visual delays.

B. Additional Mixed-Effects Model Analysis

Table 11 provides the results from the second mixed-model analysis performed to verify the effects of simulator found in the first model analysis, as introduced in Section III.F. The first model analysis investigated the effects of motion filter order O_{mf} and break frequency ω_{mf} without including the reference motion conditions FB and FM. The

Table 11 Summary of likelihood-ratio test results for mixed-effects models using simulator and motion condition as fixed effects.

	Simulator			Motion condition			Simulator × Motion Condition		
Measure	df	χ^2	p	df	χ^2	p	df	χ^2	p
RMS_e	1	5.99	0.01	7	170.36	<0.01	7	3.82	0.80
RMS_u	1	7.04	0.01	7	85.77	<0.01	7	17.55	0.01
$\sigma^2_{u_m}/\sigma^2_{u_v}$	1	3.62	0.06	6	296.54	<0.01	6	4.96	0.55
K_v	1	2.46	0.12	7	194.32	<0.01	7	11.16	0.13
K_m	1	3.41	0.06	6	85.35	<0.01	6	3.02	0.81
T_L	1	0.45	0.50	7	141.41	<0.01	7	18.82	0.01
τ_v	1	8.37	<0.01	7	6.71	0.46	7	11.68	0.11
τ_m	1	1.00	0.32	6	7.50	0.28	6	2.52	0.87
ω_{nm}	1	1.31	0.25	7	64.96	<0.01	7	8.20	0.32
ζ_{nm}	1	5.72	0.02	7	77.45	<0.01	7	2.01	0.96
$\omega_{c,t}$	1	3.14	0.08	7	111.74	<0.01	7	11.14	0.13
$\omega_{c,d}$	1	2.14	0.14	7	203.19	<0.01	7	8.90	0.26
$\varphi_{m,t}$	1	5.21	0.02	7	296.45	<0.01	7	4.53	0.72
$\varphi_{m,d}$	1	6.81	0.01	7	100.50	<0.01	7	5.48	0.60

= significant ($p < 0.05$) = not significant ($p \geq 0.05$)

second analysis includes these reference conditions. Note that no data were available for $\sigma^2_{u_m}/\sigma^2_{u_v}$, K_m, and τ_m for the no-motion condition FB, as these parameters were not estimated in this condition (hence $df = 6$ for these parameters in Table 11).

The results of the second model analysis confirm the results of the first with very minor differences. For $\sigma^2_{u_m}/\sigma^2_{u_v}$, the main effect of simulator is not significant with the reference motion conditions included (Table 11), while it is significant without these conditions included (Table 5). The second model analysis reveals an interaction between simulator and motion condition for RMS_u that is not present the first model analysis (Table 4). This is the result of the larger reduction in control intensity in the no-motion condition FB in the SRS compared to the VMS (Fig. 5(b)). The interactions of simulator with O_{mf} and ω_{mf} for K_v, ω_{nm}, and $\varphi_{m,d}$ found in the first analysis (Table 5 – Table 7) were not found in the second analysis.

References

[1] Bürki-Cohen, J., and Go, T. H., "The Effect of Simulator Motion Cues on Initial Training of Airline Pilots," *Proceedings of the AIAA Modeling and Simulation Technologies Conference and Exhibit, San Francisco (CA)*, 2005. doi:10.2514/6.2005-6109.

[2] Bürki-Cohen, J., Sparko, A. L., and Bellman, M., "Flight Simulator Motion Literature Pertinent to Airline-Pilot Recurrent Training and Evaluation," *Proceedings of the AIAA Modeling and Simulation Technologies Conference, Portland (OR)*, 2011.

[3] Schroeder, J. A., and Grant, P. R., "Pilot Behavioral Observations in Motion Flight Simulation," *Proceedings of the AIAA Guidance, Navigation, and Control Conference, Toronto, Canada*, 2010. doi:10.2514/6.2010-8353.

[4] de Winter, J. C. F., Dodou, D., and Mulder, M., "Training Effectiveness of Whole Body Flight Simulator Motion: A Comprehensive Meta-Analysis," *The International Journal of Aviation Psychology*, Vol. 22, No. 2, 2012, pp. 164–183. doi:10.1080/10508414.2012.663247.

[5] Zaal, P. M. T., and Mobertz, X. R. I., "Effects of Motion Cues on the Training of Multi-Axis Manual Control Skills," *Proceedings of the AIAA Modeling and Simulation Technologies Conference, Denver (CO)*, 2017. doi:10.2514/6.2017-3473.

[6] Zaal, P. M. T., Schroeder, J. A., and Chung, W. W. Y., "Transfer of Training on the Vertical Motion Simulator," *Journal of Aircraft*, Vol. 52, No. 6, 2015, pp. 1971–1984. doi:10.2514/1.C033115.

[7] Pool, D. M., Harder, G. A., and van Paassen, M. M., "Effects of Simulator Motion Feedback on Training of Skill-Based Control Behavior," *Journal of Guidance, Control, and Dynamics*, Vol. 39, No. 4, 2016, pp. 889–902. doi:10.2514/1.G001603.

[8] ICAO, "Manual of Criteria for the Qualification of Flight Simulation Training Devices. Volume 1 – Airplanes," ICAO Doc 9625, International Civil Aviation Organization, 2009. Third edition.

[9] FAA, "Stall Prevention and Recovery Training," Advisory Circular 120-109A, U.S. Department of Transportation, Federal Aviation Administration, Nov. 2015.

[10] Zaal, P. M. T., Popovici, A., and Zavala, M. A., "Effects of False Tilt Cues on the Training of Manual Roll Control Skills," *Proceedings of the AIAA Modeling and Simulation Technologies Conference, Kissimmee (FL)*, 2015. doi:10.2514/6.2015-0655.

[11] Popovici, A., Zaal, P. M. T., and Pieters, M. A., "Time-Varying Manual Control Identification in a Stall Recovery Task under Different Simulator Motion Conditions," *Proceedings of the AIAA Modeling and Simulation Technologies Conference, Atlanta (GA)*, 2018. doi:10.2514/6.2018-2936.

[12] Schmidt, S. F., and Conrad, B., "Motion Drive Signals for Piloted Flight Simulators," Tech. Rep. NASA CR-1601, National Aeronautics and Space Administration, Ames Research Center, 1970.

[13] Reid, L. D., and Nahon, M. A., "Flight Simulation Motion-Base Drive Algorithms. Part 1: Developing and Testing the Equations," Tech. Rep. UTIAS 296, University of Toronto, Institute for Aerospace Studies, Dec. 1985.

[14] Gouverneur, B., Mulder, J. A., van Paassen, M. M., Stroosma, O., and Field, E. J., "Optimisation of the SIMONA Research Simulator's Motion Filter Settings for Handling Qualities Experiments," *Proceedings of the AIAA Modeling and Simulation Technologies Conference and Exhibit, Austin (TX)*, American Institute of Aeronautics and Astronautics, 2003.

[15] Pool, D. M., Damveld, H. J., van Paassen, M. M., and Mulder, M., "Tuning Models of Pilot Tracking Behavior for a Specific Simulator Motion Cueing Setting," *Proc. of the AIAA Modeling and Simulation Conference, Portland (OR)*, 2011.

[16] Pool, D. M., van Paassen, M. M., and Mulder, M., "Effects of Motion Filter Gain and Break Frequency Variations on Pilot Roll Tracking Behavior," *Proceedings of the AIAA Modeling and Simulation Technologies Conference, Boston (MA)*, 2013.

[17] Pool, 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, Minneapolis (MN)*, 2012.

[18] Pool, D. M., Zaal, P. M. T., van Paassen, M. M., and Mulder, M., "Effects of Heave Washout Settings in Aircraft Pitch Disturbance Rejection," *Journal of Guidance, Control, and Dynamics*, Vol. 33, No. 1, 2010, pp. 29–41. doi:10.2514/1.46351.

[19] Van Wieringen, A. T., Pool, D. M., van Paassen, M. M., and Mulder, M., "Effects of Heave Washout Filtering on Motion Fidelity and Pilot Control Behavior for a Large Commercial Airliner," *Proceedings of the AIAA Modeling and Simulation Technologies Conference, Portland, Oregon, Aug. 8-11*, 2011.

[20] Zaal, P. M. T., Pool, D. M., Mulder, M., Van Paassen, M. M., and Mulder, J. A., "Identification of Multimodal Pilot Control Behavior in Real Flight," *Journal of Guidance, Control & Dynamics*, Vol. 33, No. 5, 2010, pp. 1527–1538.

[21] Steurs, M., Mulder, M., and van Paassen, M. M., "A Cybernetic Approach to Assess Flight Simulator Fidelity," *Proceedings of the AIAA Modelling and Simulation Technologies Conference and Exhibit, Providence (RI)*, 2004.

[22] Bergeron, H. P., "Investigation of Motion Requirements in Compensatory Control Tasks," *IEEE Transactions on Man-Machine Systems*, Vol. MMS-11, No. 2, 1970, pp. 123–125. doi:10.1109/TMMS.1970.299992.

[23] van Gool, M. F. C., "Influence of Motion Washout Filters on Pilot Tracking Performance," *Piloted Aircraft Environment Simulation Techniques*, 1978, pp. 19–1 – 19–5.

[24] Bray, R. S., "Visual and Motion Cueing in Helicopter Simulation," Technical Memorandum NASA-TM-86818, NASA Ames Research Center, Moffett Field (CA), Sep. 1985.

[25] Grant, P. R., and Reid, L. D., "Motion Washout Filter Tuning: Rules and Requirements," *Journal of Aircraft*, Vol. 34, No. 2, 1997, pp. 145–151. doi:10.2514/2.2158.

[26] Hosman, R. J. A. W., and Advani, S. K., "Are Criteria for Motion Cueing and Time Delays Possible? Part 2." *Proceedings of the AIAA Modeling and Simulation Technologies Conference, Boston (MA)*, 2013.

[27] Hodge, S. J., Perfect, P., Padfield, G. D., and White, M. D., "Optimising the roll-sway motion cues available from a short stroke hexapod motion platform," *The Aeronautical Journal*, Vol. 119, No. 1211, 2015, pp. 23–44.

[28] Correia Grácio, B. J., Valente Pais, A. R., Van Paassen, M. M., Mulder, M., Kelly, L. C., and Houck, J. A., "Optimal and Coherence Zone Comparison Within and Between Flight Simulators," *Journal of Aircraft*, Vol. 50, No. 2, 2013, pp. 493–507.

[29] Grant, P. R., Yam, B., Hosman, R. J. A. W., and Schroeder, J. A., "Effect of Simulator Motion on Pilot Behavior and Perception," *Journal of Aircraft*, Vol. 43, No. 6, 2006, pp. 1914–1924. doi:10.2514/1.21900.

[30] Hosman, R. J. A. W., Grant, P. R., and Schroeder, J. A., "Pre and Post Pilot Model Analyis Compared to Experimental Simulator Results," *AIAA Modeling and Simulation Technologies Conference and Exhibit, San Francisco (CA)*, 2005.

[31] Ellerbroek, J., Stroosma, O., Mulder, M., and van Paassen, M. M., "Role Identification of Yaw and Sway Motion in Helicopter Yaw Control Tasks," *Journal of Aircraft*, Vol. 45, No. 4, 2008, pp. 1275–1289. doi:10.2514/1.34513.

[32] Schroeder, J. A., "Helicopter Flight Simulation Motion Platform Requirements," Tech. Rep. NASA-TP-1999-208766, National Aeronautics and Space Administration, Jul. 1999.

[33] Jex, H. R., Magdaleno, R. E., and Junker, A. M., "Roll Tracking Effects of G-vector Tilt and Various Types of Motion Washout," *Proceedings of the Fourteenth Annual Conference on Manual Control*, 1978, pp. 463–502.

[34] Jex, H. R., Jewell, W. F., and Magdaleno, R. E., "Effects of Various Lateral-Beam-Motion Washouts on Pilot Tracking and Opinion in the "Lamar" Simulator," *Proceedings of the Fifteenth Annual Conference on Manual Control*, Wright State University, Dayton (OH), 1979, pp. 244–266.

[35] Zaal, P. M. T., and Zavala, M. A., "Effects of Different Heave Motion Components on Pilot Pitch Control Behavior," *Proceedings of the AIAA Modeling and Simulation Technologies Conference, Washington, D.C.*, 2016. doi:10.2514/6.2016-3371.

[36] Bates, D., Mächler, M., Bolker, B., and Walker, S., "Fitting Linear Mixed-Effects Models Using lme4," *Journal of Statistical Software*, Vol. 67, No. 1, 2015. doi:10.18637/jss.v067.i01.

[37] van 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, Kyoto Japan*, 1998, pp. 515–520.

[38] Zaal, P. M. T., Pool, D. M., Chu, Q. P., van Paassen, M. M., Mulder, 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. doi:10.2514/1.42843.

[39] Shirley, R. S., and Young, L. R., "Motion Cues in Man-Vehicle Control – Effects of Roll-Motion Cues on Human Operator's Behavior in Compensatory Systems with Disturbance Inputs," *IEEE Transactions on Man-Machine Systems*, Vol. 9, No. 4, 1968, pp. 121–128. doi:10.1109/TMMS.1968.300016.

[40] Zaal, P. M. T., Pool, D. M., de Bruin, J., Mulder, M., and van Paassen, M. M., "Use of Pitch and Heave Motion Cues in a Pitch Control Task," *Journal of Guidance, Control, and Dynamics*, Vol. 32, No. 2, 2009, pp. 366–377. doi:10.2514/1.39953.

[41] McRuer, D. T., and Jex, H. R., "A Review of Quasi-Linear Pilot Models," *IEEE Transactions on Human Factors in Electronics*, Vol. HFE-8, No. 3, 1967, pp. 231–249. doi:10.1109/THFE.1967.234304.

[42] McRuer, D. T., Magdaleno, R. E., and Moore, G. P., "A Neuromuscular System Activation Model," *IEEE Transactions on Man-Machine Systems*, Vol. MMS-9, No. 3, 1968, pp. 61–71.

[43] McRuer, D. T., Graham, D., Krendel, E. S., and Reisener, W. J., "Human Pilot Dynamics in Compensatory Systems: Theory, Models, and Experiments with Controlled Element and Forcing Function Variations," Tech. Rep. AFFDL-TR-65-15, Air Force Flight Dynamics Laboratory, Wright-Patterson Air Force Base (OH), 1965.

[44] Sinacori, J. B., "The Determination of Some Requirements for a Helicopter Research Simulation Facility," Tech. Rep. NASA-CR-152066, Systems Technology Inc., Sep. 1977.

[45] Box, G. E. P., and Cox, D. R., "An Analysis of Transformations," *Journal of the Royal Statistical Society*, Vol. 26, No. 2, 1964, pp. 211–252.

[46] Valente Pais, A. R., Pool, D. M., de Vroome, A. M., van Paassen, M. M., and Mulder, M., "Pitch Motion Perception Thresholds During Passive and Active Tasks," *Journal of Guidance, Control, and Dynamics*, Vol. 35, No. 3, 2012, pp. 904–918. doi:10.2514/1.54987.

[47] Wijlens, R., Zaal, P. M. T., and Pool, D. M., "Retention of Manual Control Skills in Multi-Axis Tracking Tasks," *Proceedings of the AIAA Modeling and Simulation Technologies Conference, Orlando (FL)*, 2020. doi:10.2514/6.2020-2264.

[48] Hosman, R. J. A. W., and van der Vaart, J. C., "Effects of Vestibular and Visual Motion Perception on Task Performance," *Acta Psychologica*, Vol. 48, 1981, pp. 271–287.

[49] Pool, D. M., Mulder, M., van Paassen, M. M., and van der Vaart, J. C., "Effects of Peripheral Visual and Physical Motion Cues in Roll-Axis Tracking Tasks," *Journal of Guidance, Control, and Dynamics*, Vol. 31, No. 6, 2008, pp. 1608–1622. doi:10.2514/1.36334.

[50] Peterse, H. P. M., Pool, D. M., van Paassen, M. M., and Mulder, M., "Interactions of Outside Visual Cues and Motion Cueing Settings in Yaw Tracking," *Proceedings of the AIAA Modeling and Simulation Technologies Conference, Washington, D.C.*, 2016. doi:10.2514/6.2016-3370.

[51] Mendes, M. S. F., Pool, D. M., and van Paassen, M. M., "Effects of Peripheral Visual Cues in Simulator-Based Training of Multimodal Control Skills," *Proceedings of the AIAA Modeling and Simulation Technologies Conference, Denver (CO)*, 2017. doi:10.2514/6.2017-3671.

[52] Aponso, B. L., Beard, S. D., and Schroeder, J. A., "The NASA Ames Vertical Motion Simulator – A Facility Engineered for Realism," *Proceedings of the Royal Aeronautical Society Spring 2009 Flight Simulation Conference, London, UK*, 2009.

[53] Beard, S. D., Reardon, S. E., Tobias, E. L., and Aponso, B. L., "Simulation System Optimization for Rotorcraft Research on the Vertical Motion Simulator," *Proceedings of the AIAA Modeling and Simulation Technologies Conference, Minneapolis (MN)*, 2012. doi:10.2514/6.2012-4634.

[54] Nieuwenhuizen, F. M., Mulder, M., van Paassen, M. M., and Bülthoff, H. H., "Influences of Simulator Motion System Characteristics on Pilot Control Behavior," *Journal of Guidance, Control, and Dynamics*, 2013. doi:10.2514/1.59257.

[55] Berkouwer, W. R., Stroosma, O., van Paassen, M. M., Mulder, M., and Mulder, J. A., "Measuring the Performance of the SIMONA Research Simulator's Motion System," *Proceedings of the AIAA Modeling and Simulation Technologies Conference and Exhibit, San Francisco (CA)*, 2005. doi:10.2514/6.2005-6504.

[56] Lehmer, R. D., and Chung, W. W. Y., "Image Dynamic Measurement System (IDMS-2) for flight simulation fidelity verification," *Modeling and Simulation Technologies Conference and Exhibit*, AIAA, 1999. doi:10.2514/6.1999-4035.

[57] Stroosma, O., van Paassen, M. M., Mulder, M., and Postema, F. N., "Measuring Time Delays in Simulator Displays," *Proceedings of the AIAA Modeling and Simulation Technologies Conference and Exhibit, Hilton Head (SC)*, 2007.