EFFECTIVENESS OF A COMPUTER-BASED HELICOPTER TRAINER FOR INITIAL HOVER TRAINING

Paolo F. Scaramuzzino[1,2], Giulia D'Intino[3,4], Stefano Geluardi[3], Marilena D. Pavel[1], Daan M. Pool[1],
Olaf Stroosma[1], Max Mulder[1], and Heinrich H. Bülthoff[3]

[1]Delft University of Technology (The Netherlands)
{p.f.scaramuzzino, m.d.pavel, d.m.pool, o.stroosma, m.mulder}@tudelft.nl

[2]Politecnico di Milano (Italy)

[3]Max Planck Institute for Biological Cybernetics (Germany)
{giulia.dintino, stefano.geluardi, heinrich.buelthoff}@tuebingen.mpg.de

[4]University of Pisa (Italy)

Abstract

Today, simulators are achieving levels of complexity and cost that are comparable to those of the aircraft they should replace. For this reason, questions have been raised, in both the technical and training communities, on the required level of simulation fidelity for effective pilot training. Computer Based Trainers (CBTs) are not currently considered in regulatory standards, because it has not been proven yet whether they can replace or complement actual flight training hours. The aim of this paper is to better understand to what extent the low-level hover skills developed on a CBT are effectively transferred to a more realistic simulation environment. To achieve this goal, a quasi-Transfer-of-Training (qToT) experiment with task-naïve participants was performed in the CyberMotion Simulator (CMS) at the Max Planck Institute for Biological Cybernetics. Twenty-four subjects, divided in two groups, were trained to perform the hover maneuver controlling an identified model of a Robinson R44 civil light helicopter. The first group (the "*experimental*" group) was trained in a CBT and then transferred to the realistic setting in the CMS. The second group (the "*control*" group) received the entire training in the CMS. At the end of the experiment, the two groups were found to show comparable performance. This suggests that, even for the training of low-level flying skills, CBTs may be a valid alternative to high fidelity simulators, if supported by a suitable training program.

1. INTRODUCTION

Flight Simulator Training Devices (FSTDs) are crucial tools for pilot training. These devices are cost effective, flexible, and provide an inherently safe environment for training even hazardous scenarios[1]. Simulators have been used for over a century to aid trainees in the acquisition, development, and maintenance of their flying skills without leaving the ground[2–4]. For rotorcraft it is highly desirable to be able to develop low-level flying skills in simulators, given the intrinsically difficult helicopter flight dynamics.

Since the computer software and hardware incorporated into a FSTD determine its developmental, operational, and maintenance costs, there is great academic and industrial interest in understanding simulation fidelity requirements needed to meet FSTDs users' needs[5]. One flexible and affordable training solutions suitable for novice pilots currently considered is the low-fidelity "Desktop Trainer", also known as Computer Based Trainer (CBT).

Especially for training that makes use of low-fidelity CBT, it is critical to experimentally prove the effectiveness of the supplied training and the transfer of learned skills to the real world setting. Transfer-of-Training (ToT) experiments are one of the few available techniques that can be used to

---

explicitly measure such training effectiveness. Numerous studies have been dedicated to verifying the effectiveness of training in CBT. Unfortunately, many investigations focus on instrument[6,7] and situation awareness[8] training only. Furthermore, in those studies that explicitly investigated the training of flying skills in CBT, the experimental evidence for training effectiveness has not always been consistent. For example, Ortiz[9] trained sixty college students with no previous flight experience to perform a squared pattern maneuver. In this case even a true ToT experiment design was used: thirty of the subjects were trained in a CBT before flying the actual aircraft, while the remaining thirty received real-flight training only (Cessna 150 and 152). Statistical tests on the measured data showed that in the real aircraft the CBT-trained experimental group performed significantly better than the control group. In a separate study, Proctor et al.[10] considered three different interface configurations (cabin with motion, cabin with no motion and CBT) and trained participants to perform a complex task of combat search and rescue, while controlling a model of the UH-60. Although not being a ToT experiment, their results showed that learning did not occur in the helicopter Computer Based Trainer, arguing that the provided time frame to master the task might not be acceptable to many possible users because of the monitor size. Recent investigations by Fabbroni et al.[11,12], however, showed that hover skills acquired during fixed-base training in a CBT with a wide field-of-view display do transfer to a more realistic setting in a full-motion flight simulator.

The goal of this paper is to explicitly evaluate the extent to which hover skills developed on a Computer Based Trainer are effectively transferred to a more realistic environment. To achieve this goal, a quasi-Transfer-of-Training (qToT) experiment[13] with task-naïve participants was performed. In this experiment, the moving-base CyberMotion Simulator (CMS) at the Max Planck Institute for Biological Cybernetics, shown in Fig. 1, was used as the transfer environment.

The paper is structured as follows. Section 2 describes the experimental design and set-ups that were used. In Section 3 the results of the experiment are presented. The results are discussed and conclusions are drawn at the end of the paper.

2. EXPERIMENT DESIGN

In the experiment, participants with no prior flight experience neither in actual helicopters nor in simulators, were trained to perform the hover maneuver controlling an identified model of a Robinson R44 civil light helicopter[15,16]. Two groups were considered. The first group (the "*experimental*" group) was trained on a CBT (Fig. 2) and then transferred to the CMS. The second group (the "*control*" group) received the entire training in the CMS. A previously developed hover training program[11] was used to bring participants to a satisfactory level of performance. Previous works proved the effectiveness of the adopted training, which is designed as a realistic flight lesson divided into phases.

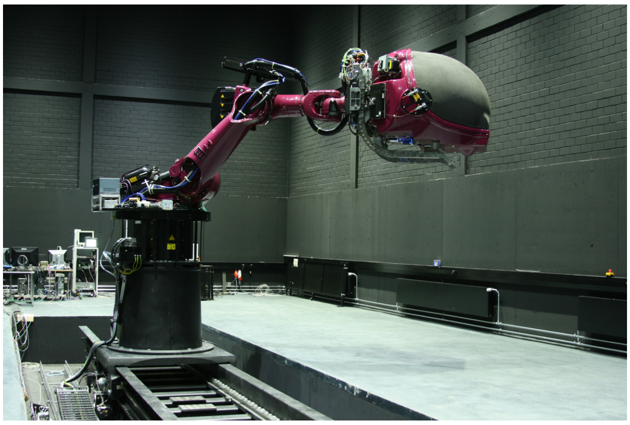


Figure 1: The MPI CyberMotion Simulator[14].

Figure 2: The Computer Based Trainer set-up used in this experiment, equipped with the Pro Flight Trainer© PUMA helicopter control inceptor.

2.1. Participants

A total of twenty-four task-naïve participants took part in the experiment, fifteen male and nine female. The participants had an average age of 26 years ($\sigma = \pm 3.81$ years). They were chosen based on a pre-experimental aptitude test intended to select for good manual control skills.

2.2. Aptitude Test

An effort was made to select the participants and to balance them equally over the two groups based on the performance they achieved in a two degrees-of-freedom (pitch and roll) combined target-following and disturbance-rejection task. The task consisted of rejecting a disturbance signal acting on the controlled element, i.e., the dynamics that each subject had to control. The controlled element had dynamics that resemble the dynamics of an aircraft and for both the pitch and roll axes were described by the following transfer function:

$$H_{ce} = \frac{K_{ce}}{s\,(s+1)} = \frac{1}{s\,(s+1)} \tag{1}$$

The results of this test and the procudure followed to form the two groups are summarized in Appendix A.

2.3. Experiment Structure

The main experiment was divided into three phases (Familiarization, Training and Evaluation) and was carried out on two different days, as shown in Tab. 1. In total, each participant was trained in the simulators for approximately 3 hours.

2.3.1. Familiarization

All participants were asked to read a short briefing document before starting the experiment, explaining the basic setup of the experiment and the task to be performed. Besides the general information concerning the experiment setup, a detailed instruction was provided regarding the helicopter dynamics and flight controls, the presented visual environment and the program intended to teach the execution of the hover maneuver through a step-by-step training. This training program consisted of five tasks of increasing level ofdifficulty , summarized in Tab. 2. These tasks were selected based on consultations with a helicopter instructor pilot (IP) and based on the results of previous training experiments[11,12,17]. Specifically, they were defined as follows:

1. *Left/Right Hovering Turn*. In this task, participants control only the pedals. All the other axes are controlled by the autopilot described in Appendix B. This maneuver starts in a stabilized hover at an altitude of 25 ft (≈ 7.5 m) in front of a hover board, placed 360 ft (≈ 110 m) in front of the starting position (see Fig. 3). The target is oriented 90° to the left and identified by an equally distant hover board. After reaching the target, the heading is to be maintained for 10 seconds. This maneuver is then repeated for a target oriented 90° to the right.
2. *Up/Down Vertical Repositioning*. In this task, participants control only the collective. All the other axes are controlled by the autopilot described in Appendix B. This maneuver starts in a stabilized hover at an altitude of 25 ft (≈ 7.5 m) in front of a hover board, placed 300 ft (≈ 90 m) in front of the starting position. Additionally, a blue sphere is placed halfway between the starting position and the hover board to aid the participant in maintaining the correct vertical position. The target is placed 50 ft (≈ 15.25 m) above the starting position and identified by an equally distant hover board. After reaching the target, the altitude is to be maintained for 10 seconds. This maneuver is then repeated in the opposite direction, starting in a stabilized hover at an altitude of 75 ft (≈ 22.75 m).
3. *Up/Down Vertical Repositioning and Heading Hold*. This maneuver is analogous to the previous one, except for the fact that the participants also control the pedals and have to compensate for the couplings related to the use of the collective lever.
4. *Hover A*. In this task, participants control only the cyclic. All the other axes are controlled by the autopilot described in Appendix B. This maneuver starts in a stabilized hover at an altitude of 25 ft (≈ 7.5 m) in front of a hover board, placed 360 ft (≈ 110 m) in front of the starting position. The participants objective is to maintain the helicopter in hover for 30 s minimizing position and heading error with respect to the initial position.
5. *Hover B*. This maneuver is analogous to the previous one. However, in this case the participants also control the pedals and the collective.

2.3.2. Training

During the experiment's Training phase (see Tab. 1), participants were asked to perform the Hover B maneuver for 30 trials of 30 seconds each in the simulator assigned to their group (CBT or CMS). During the first three trials of the Training phase, the CMS motion was disabled in order help participants of the CMS group get acquainted with the unaugmented helicopter. Hence, these trials were neglected.

2.3.3. Evaluation/Transfer

After training, the experimental group (CBT) was transferred to the CyberMotion Simulator (CMS). Participants of both groups were asked to perform again the Hover B maneuver for 30 trials of 30 seconds each. During the first three trials of the Evaluation phase, participants of the CBT group were trained in the CMS without motion in order to get acquainted with the new simulation environment. Hence, these trials were neglected.

2.4. Hypothesis

The participants of the CBT group performed the Training phase relying solely on the visual cues produced by a 22.5 in desktop monitor. During this phase, their visual sensory system adapts to the small screen size. It is expected that training in simulation environments with poor cues will enhance perceptual learning. The improved perception skills of the participants of the CBT group can allow them to adjust their control strategy in order to adapt to the available cues in the new simulator. Thus, it is expected that the hover performance of the CBT group won't be worse than that achieved by the control group, once transferred to the CMS.

2.5. Independent Variables

The qToT experiment described in this paper is influenced by three main distinct features of the two considered simulators:

- The presence of *motion cues*. This feature is crucial to evaluate the transfer of training from a fixed-base to a moving-base simulator;
- The *display type*. This feature influences the transfer from a desktop monitor to a large FOV cabin equipped with two projectors;
- The *immersiveness* of the simulation, determined by the difference between an office desk and the CMS cabin.

Because of the impossibility to isolate the individual contribution of each feature to the transfer of training, only one independent variable was in fact considered, i.e., the overall simulator's fidelity.

2.6. Dependent Variables

To investigate the effect of simulator's fidelity (independent variable) on hover performance, the following dependent measures were defined:

- *Number of completed trials*. The number of trials in which the control of the helicopter model was not lost for the full duration of the trial. This index can be used as an indication of the training effectiveness in maneuvers where the stability of the helicopter is not guaranteed by the controller, as in Hover A and Hover B (Table 2).
- *Position Scores*. The root mean squared (RMS) position error with respect to the target hover position was calculated at the end of each completed trial for longitudinal (x), lateral (y) and vertical (z) positioning and for the position magnitude ($P = \sqrt{x^2 + y^2 + z^2}$). Eq. (2) shows how these metrics are calculated, taking the longitudinal positioning as example.

  $$\text{RMS}_x = \sqrt{\frac{1}{N}\sum_{k=1}^{N}\left[x(k) - x(1)\right]^2} \quad (2)$$

  where N is the number of time samples considered in the trial.

  These indexes can be used to objectively evaluate the student pilots' performance while executing the maneuvers.
- *Heading Score*. The root mean squared (RMS) heading error was calculated at the end of each completed trial.
- *Velocity Score*. The root mean square (RMS) of the linear velocity was calculated at the end of each completed trial.

  This index can be used as an indication of hover stability.
- *Control activity*. To gain insights into the participants' control activity, the root mean squared (RMS) deviation with respect to the trim position for every helicopter control was computed at the end of each completed trial.

The part-task training during the Familiarization phase was time-based. Therefore, the total number of trials performed in each task is different for each participant. For this reason, results presented in this paper focus on Training and Evaluation phases only.

Table 1: Experiment phases.

Phase	Experimental group	Control group	Duration
Familiarization (Day 1)	Instructions session	Instructions session	15 minutes
	Part-task training in the CBT (Tab. 2)	Part-task training in the CMS (Tab. 2)	1 hour and 45 minutes
Training (Day 1)	Hover with all controls in the CBT	Hover with all controls in the CMS	30 trials of 30 seconds each
Evaluation/Transfer (Day 2)	Hover with all controls in the CMS	Hover with all controls in the CMS	30 trials of 30 seconds each

Table 2: Part-task training tasks.

m_{ID}	Task	Controls used	Duration
1	Left/right Hovering Turn	Pedals	5 min
2	Up/down Vertical Repositioning	Collective	5 min
3	Up/down Vertical Repositioning, Heading Hold	Collective + Pedals	20 min
4	Hover A	Cyclic	30 min
0	Hover B	Cyclic + Collective + Pedals	30 min

2.7. Apparatus

This Section provides a description of the two considered helicopter simulators.

The CyberMotion Simulator (CMS) in Fig. 1 is an anthropomorphic robotic arm (KUKA Roboter, GmbH) mounted on a linear rail to provide a total of 8 degrees-of-freedom. Thanks to its high agility and motion envelope, the CMS is well suited for helicopter hover training. The end-effector consists of a custom-built helicopter cockpit with a 140° horizontal for 70° vertical field-of-view that allows for virtual environments to be projected. For the experiment described in this paper, the cockpit was also equipped with a pilot seat and a commercial off-the-shelf helicopter control inceptor (Pro Flight Trainer PUMA) with no programmable control loading systems.

The motion of the CMS was generated by means of a classical Motion Cueing Algorithm (MCA) based on second-order high-pass washout filters[18,19]. The gains were manually tuned based on the evaluations of four expert Robinson R44 pilots, until a good matching between visual and motion cues was achieved.

The Computer Based Trainer (CBT) in Fig. 2 is equipped with a pilot seat, a 22.5 in display and the same control inceptor used in the CyberMotion Simulator (CMS). The display is produced by VIEWPixx, VPixx Technologies Inc., Canada.

As discussed at the beginning of the paper, in this experiment, an identified model of a Robinson R44 civil light helicopter was used. This model was developed in previous research and experimentally validated[16].

The visual environment projected in the two simulators was developed in Unity®[20], see Fig. 3. It displays the inside of a Robinson R44 cockpit, while the out-of-the-window scenery consists of a heliport with a wide field in which the helicopter can move without encountering any obstacle. Markers, such as lines and dots, were drawn on the heliport ground to help the participants understand position and attitude of the helicopter. Moreover, hover boards were placed in the scenery and were used by the student pilots as reference points for accomplishing the experiment tasks. An artificial horizon, in the form of a head-up display, was also added to help the pilot estimate the attitude of the vehicle even for the experimental condition without motion cues (CBT).

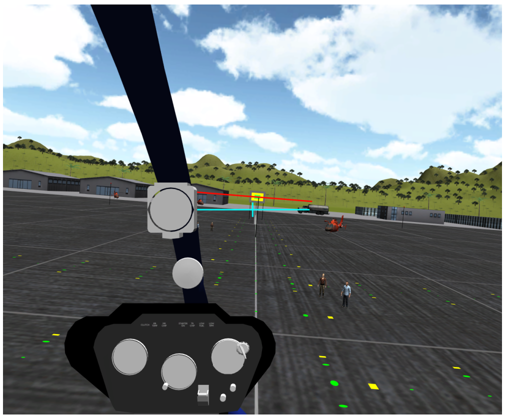


Figure 3: Hover scenario visual scene.

3. RESULTS

The experimental results will be presented in the following figures as box-whiskers plots. On each box, the white circle represents the median over different data points. The box is delimited by the first and third quartiles, therefore it includes data points between the 25th and the 75th percentile. The difference between first and third quartiles defines the interquartile range. The two edges of the whiskers indicate the lowest and the highest data point within 1.5 of the interquartile range. All the data points not included in the whiskers are considered as outliers and they are represented by cross markers. The dashed line, displayed in some of the plots, represents the y-axis upper limit. Any data value, that falls outside it, is displayed evenly distributed in the adjacent region, retaining the relative order of the points.

3.1. Completed Trials

Fig. 4 shows the absolute and relative numbers of completed trials by participants of both groups in each phase. The data points on which each box plot is based are plotted next to it (filled circle markers), together with the mean value (diamond marker). It can be noticed that the experimental group (CBT) had a higher success rate than the control group (CMS) during the training phase, with an average number of completed trials that is almost twice as high (Tab. 3). This marked difference disappears in the evaluation phase, where performance of the CBT group remains almost unchanged. In the last session of the experiment, participants of both groups were able to stabilize the helicopter model in the CMS, on average, in the 60% of the runs, suggesting the effectiveness of the training program.

The dramatically smaller number of completed runs for the CMS group during the training phase is, in hindsight, related to the stricter safety limits in the CMS. Furthermore, some of the participants in the CMS group may have been overwhelmed by the CMS, which is characterized by high vibrations level and by a small cabin equipped with a large FOV projection screen.

Table 3: Group performance comparison in terms of average number of completed trials.

Phase	Group	
	CBT	CMS
Training	19/30 (63%)	10/30 (33%)
Evaluation	17/30 (58%)	18/30 (61%)

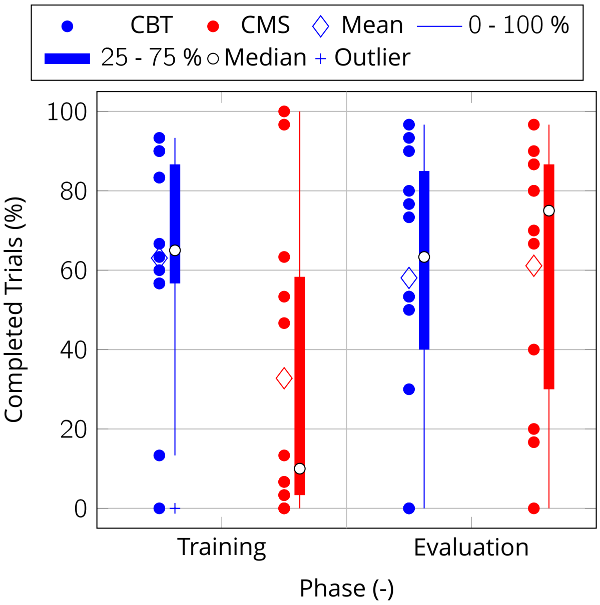


Figure 4: Distribution of the percent number of completed trials by participants of both groups in each phase.

3.2. Performance Scores

The evolution of participants' performance is shown in Fig. 5 in terms of longitudinal position. This score was found to be the most illustrative of the performance score parameters considered in this experiment. The number that appears on the top (CMS group) or at the bottom (CBT group) of each boxplot represents the number of samples available, i.e., the number of participants that completed the corresponding trial. This additional information is provided in order to avoid a misleading interpretation of the results due to differences in the number of completed trials. At first glance, focusing on the training phase (Fig. 5a), the CMS group performs better than the CBT group, exhibiting also a lower within-group variability. However, for each trial the boxplot related to the CMS group is based on a number of samples that is, on average, half of the number of samples available for the CBT group (Fig. 4).

The CMS group shows a learning trend over the first half of the training phase, followed by a degradation in the performance registered in the last part of this phase. The CBT group displays fluctuating, but overall flat performance. The first session of the experiment culminated in the training phase and lasted approximately 2 hours and 30 minutes. Hence, the data in Fig. 5a suggest that participants may have been affected by fatigue towards the end of the session.

In the Evaluation phase (Fig. 5b), the comparison between the two groups becomes fairer with respect to the Training phase. Indeed, for each trial the boxplots of the two groups are based, on average, on the same number of samples (Fig. 4). Neglecting the first three trials of the Evaluation Phase (Section 2.3.3), the CBT group almost immediately reaches performance comparable to that achieved by the CMS group, but even better in terms of within-group variability. For both groups, a learning trend appears in the second half of the evaluation phase. This trend is more pronounced for the CBT group.

The effectiveness of the training was further investigated by averaging the scores defined in Section 2.6 over the completed trials by each participant. These metrics are shown in Fig. 6 as box-whiskers plots to compare the performance of the two groups in the Training and in the Evaluation phases. Boxplots are plotted together with the data points on which they are based. Each data point corresponds to one participant and the number that appears next to it represents the number of completed trials by that participant. As can be seen in Tab. 4, the CBT group significantly improved its performance from the training phase to the evaluation phase for every considered metric, except for the vertical score and the heading score.

For some metrics (longitudinal, heading, position and velocity scores), the enhancement of the performance is associated with a decrease of the within-group variability.

No significant difference was found between the two phases for the CMS group (Tab. 4). The participants of this group were not able to stabilize the helicopter in a large number of trials during the training phase. During the evaluation phase, they reach a level of performance close to that shown by the participants who were able to complete the task throughout the training phase. The increase in the number of completed trials in the evaluation phase (Tab. 3) leads to a growth in the within-group variability for almost every performance metric.

Tab. 5 shows that the two groups achieved comparable performance. Indeed, the data of the two groups were not statistically different in any phase of the experiment. The largest difference was found for the longitudinal score during the training phase ($t(19) = 1.852$, $p = 0.08$) and is again related to the small number of trials completed by the CMS group.

Table 4: Dependent-samples T test between training phase and evaluation phase.

Metric	Group	t-test		
		t	df	Sig. (2-tailed)
$\overline{\text{RMS}}_x$	CBT	4.570	9	0.001*
	CMS	−0.949	9	0.368
$\overline{\text{RMS}}_y$	CBT	3.075	9	0.013*
	CMS	−0.022	9	0.983
$\overline{\text{RMS}}_z$	CBT	−0.813	9	0.437
	CMS	−0.816	9	0.435
$\overline{\text{RMS}}_\psi$	CBT	0.787	9	0.451
	CMS			0.508[a]
$\overline{\text{RMS}}_P$	CBT	3.826	9	0.004*
	CMS	−0.556	9	0.592
$\overline{\text{RMS}}_V$	CBT	5.462	9	0.000*
	CMS	1.058	9	0.318

* Significant ($p < 0.05$) difference between compared samples.

[a] At least one sample not normally distributed. Related-samples Wilcoxon signed-rank test was applied instead of paired-samples T test.

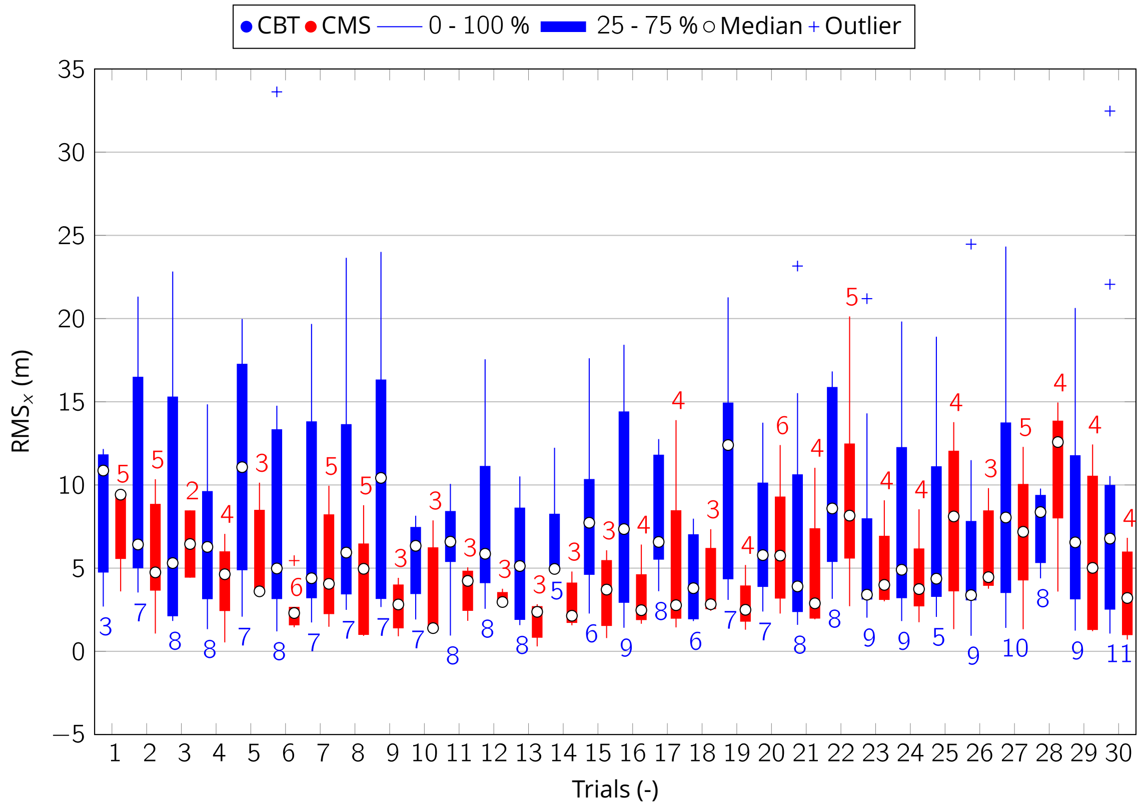
Table 5: Independent-samples T test between the two groups.

Metric	Phase	t-test		
		t	df	Sig. (2-tailed)
$\overline{\text{RMS}}_x$	T	1.852	19.000	0.080
	E	−0.930	19.000	0.364
$\overline{\text{RMS}}_y$	T	1.103	19.000	0.284
	E	−0.352	19.000	0.729
$\overline{\text{RMS}}_z$	T	−1.505	19.000	0.149
	E	−0.989	14.947	0.339
$\overline{\text{RMS}}_\psi$	T			0.251[a]
	E	0.015	19.000	0.988
$\overline{\text{RMS}}_P$	T	1.633	19.000	0.119
	E	−0.678	19.000	0.506
$\overline{\text{RMS}}_V$	T	1.004	19.000	0.328
	E	−0.638	19.000	0.531

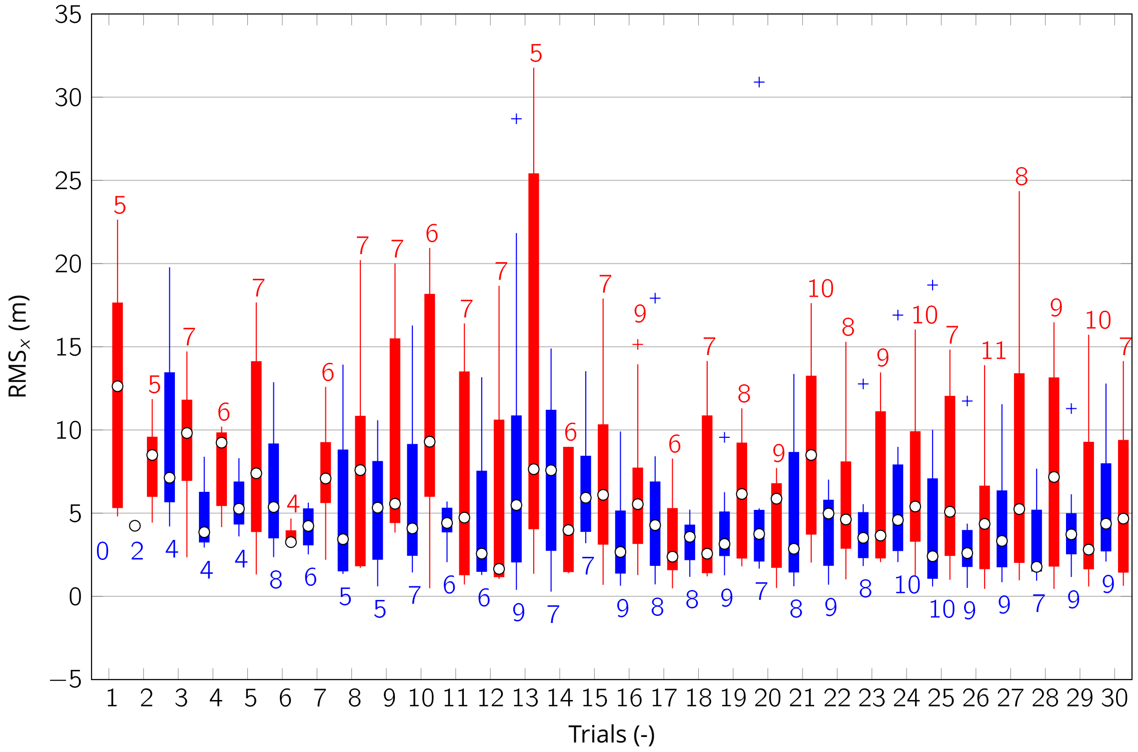
[a] At least one sample not normally distributed. Independent-samples Mann-Whitney U test was applied instead of Independent-samples T test.

3.3. Control Effort

In order to justify some of the results obtained in terms of performance, it is worth looking also at the participants' control activity. As shown in Fig. 6c and 6d, vertical and heading scores were the only two metrics in which no improvement was noticed from the training to the evaluation phase for the CBT group. This might be related to how participants were briefed. They were taught to first stabi-

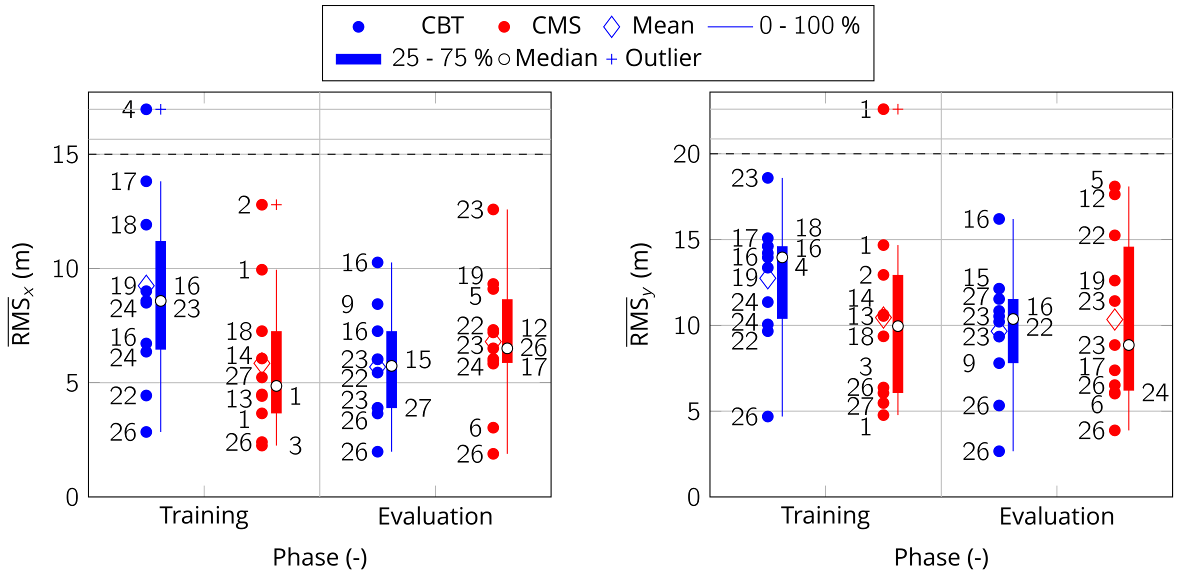


(a) Training phase.



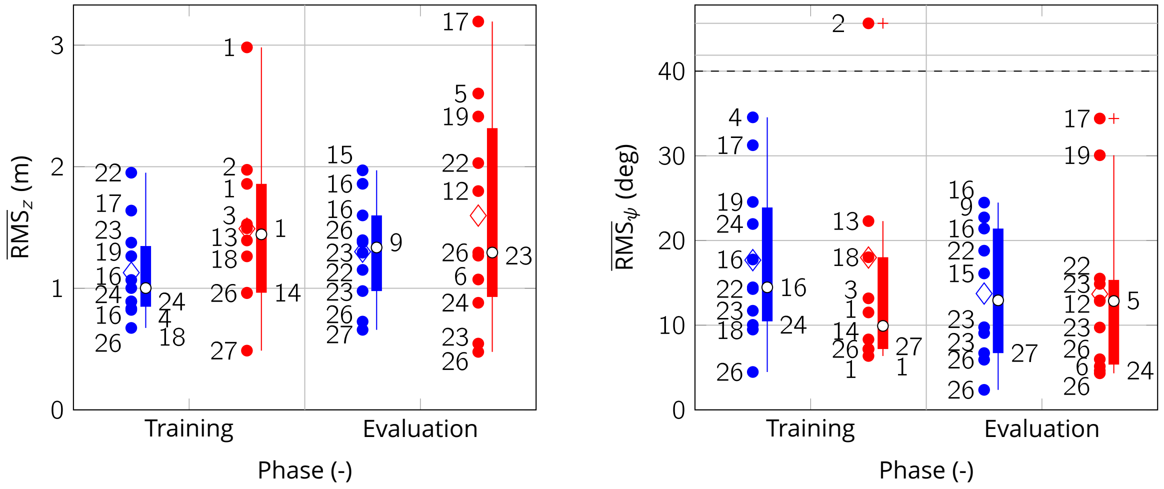
(b) Evaluation phase.

Figure 5: Evolution of the distribution of the longitudinal score for each group.



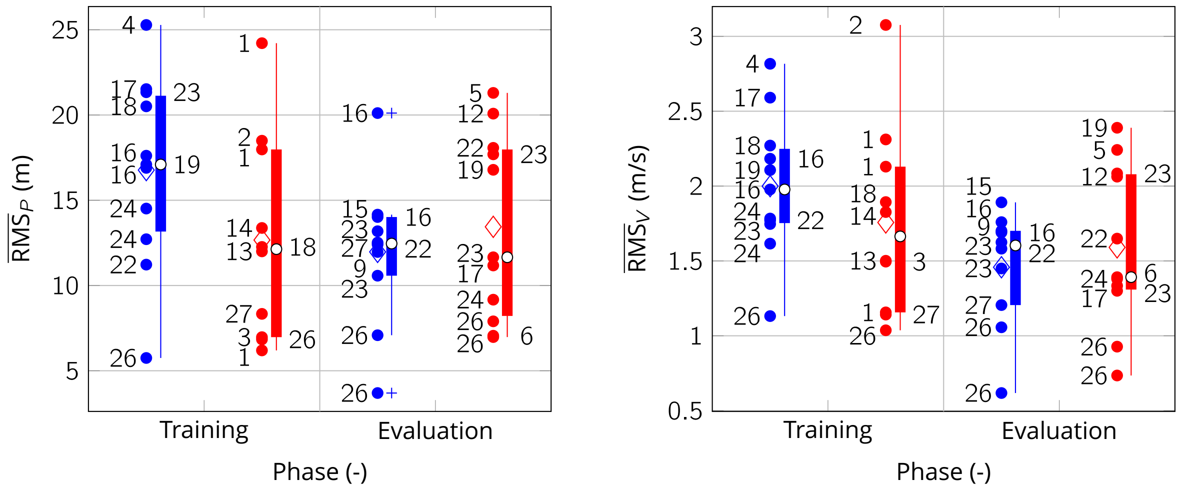
(a) Longitudinal score.

(b) Lateral score.



(c) Vertical score.

(d) Heading score.



(e) Position score.

(f) Linear velocity score.

Figure 6: Distribution of the average score for each group in each phase.

lize the helicopter, giving priority to the use of the cyclic stick. Thereafter, within the same run, they were required to hover, using the pedals and the collective lever to make adjustments of the helicopter's heading and altitude. This is also proven by the fact that both groups exhibit lower control activity for the pedals and the collective than for the longitudinal and the lateral cyclic (Fig. 7). Furthermore, for both groups there was no change in terms of control activity from the training to the evaluation phase for the longitudinal cyclic (Fig. 7a), the collective (Fig. 7c) and the pedals (Fig. 7d). Conversely, a decrease in the control activity from training to evaluation phase can be noticed for the lateral cyclic (Fig. 7b) for the CMS group, suggesting a reduction in the workload required from participants to stabilize the helicopter model. This can also be inferred from the increase in the number of trials completed by the CMS group in the evaluation phase.

4. DISCUSSION

The experiment presented in this paper was designed to investigate how effective a CBT can be for hover training of novice pilots. The results of this quasi-Transfer-of-Training experiment confirm previous results in literature which showed how the effectiveness of a Flight Simulator Training Device (FSTD) depends more on the design of the training program than on the fidelity provided by the simulator itself[21]. Indeed, after approximately 2 hours and 30 minutes of practice in the respective simulators, the two groups of participants (CBT and CMS groups) showed almost identical proficiency levels in the evaluation phase.

Helicopters are unstable in hover, but the pilot acts as a feedback controller and uses the available cues as source of information to close the loop and stabilize the system. Experienced pilots are taught to give priority to some of the available cues depending on the flight condition, but in general they are supposed to trust their instruments and ignore their vestibular sensory input. Despite this, simulator motion bases enable better in-simulator performance by experienced pilot and there is nearly unanimous preference to have this feature implemented in simulators[22]. Novice pilots, instead, apparently gather information from the visual sensory system disregarding the presence of motion cues in hover and low-speed maneuvers[17] (Tab. 5), at least for the current MCA set of parameters.

The quasi-Transfer-of-Training experiment showed that the part-task training was effective in teaching the basics of helicopter dynamics and control. Indeed, participants of both groups were able to consistently stabilize a Robinson R44 identified model at the end of the evaluation phase.

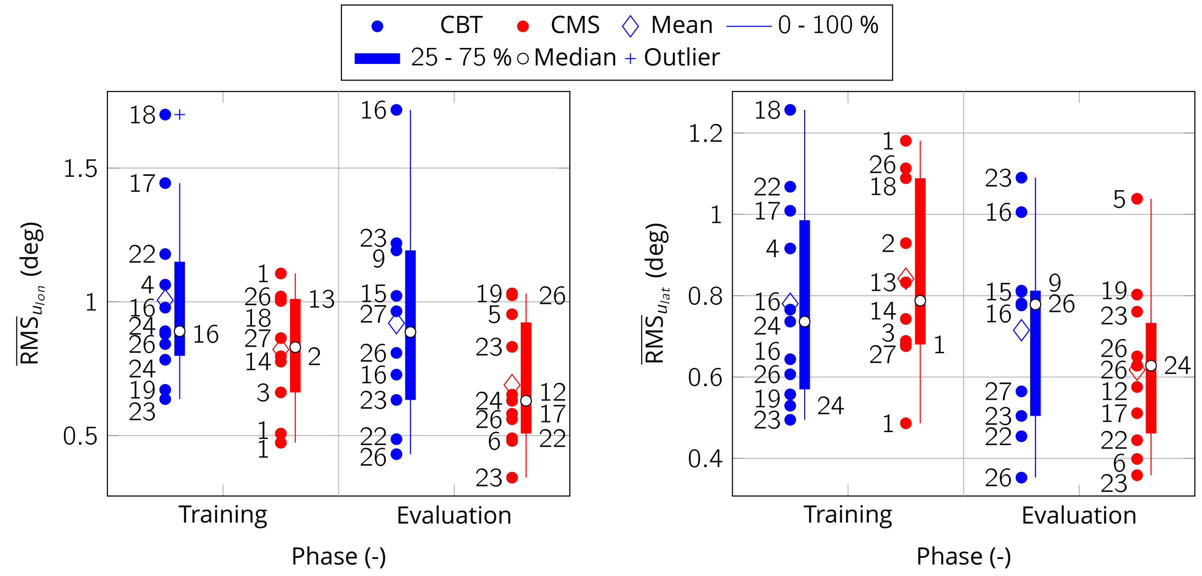
The biggest difference between the two groups occurred during the training phase and was related to the number of completed trials. In particular, the participants of the CBT group were able to complete, on average, 63% of the total number of trials for this phase, against only 33% of the CMS group. A possible reason for this result is that participants of the CMS group might have been overwhelmed by the impact with the CMS, which is characterized by high vibrations level and by a small cabin equipped with a large FOV projection screen. The duration of the first session of the experiment (2 hours and 30 minutes) might have increased the level of stress and fatigue, affecting the results of the training phase for the CMS group. As a future recommendation, it is advisable to split the experiment in three sessions in order to mitigate the influence of participants' fatigue on the results. Furthermore, biophysical measurements can be used in future studies to evaluate participants' workload and to determine if stress and fatigue were actually confounding factors.

The CBT group showed significant improvement in performance from the training phase to the evaluation phase for all the considered metrics, except for the vertical and the heading scores. As a consequence, it can be concluded that the pedals and the collective require additional attention during the part-task training, not only when they are used separately, but also in combination.

From the analysis carried out on the collected data, no differences between the CBT and CMS groups were found. Although the relatively low number of participants does not result in sufficient statistical power, the obtained results seem to confirm our hypothesis that CBTs may be a valid alternative to high-fidelity simulators in the training of task-naïve helicopter pilots, if supported by a suitable training program.

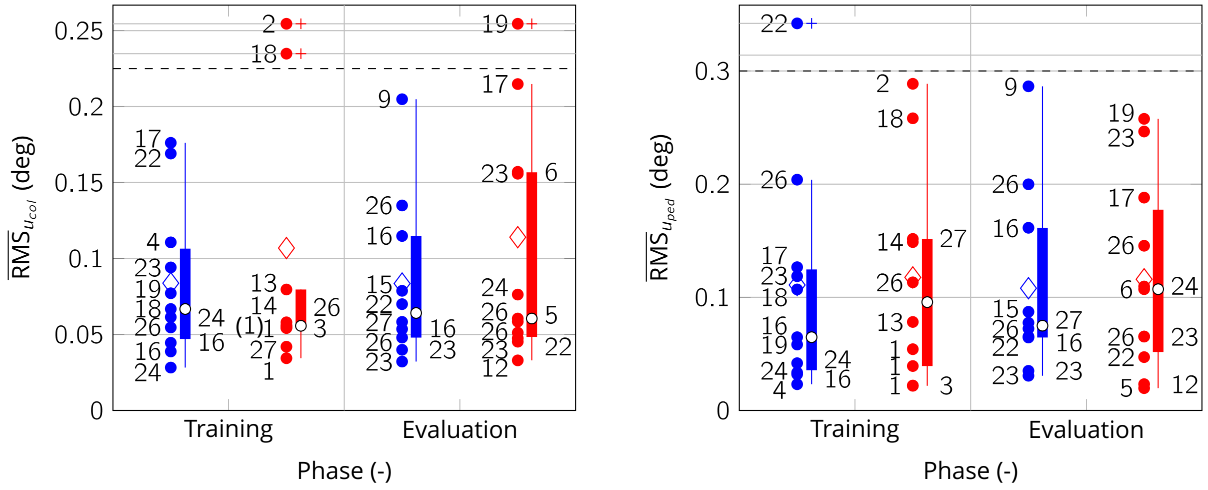
5. CONCLUSIONS

This paper presented the results of a quasi-Transfer-of-Training experiment performed to compare the effectiveness of low- and high-fidelity flight simulators to train the hover maneuver to task-naïve helicopter pilots. Participants were divided into two groups: one trained in a Computer Based Trainer and one in the MPI CyberMotion Simulator. The training session was followed by an evaluation session in which the group trained in the CBT was transferred to the CMS to evaluate the effects of the simulator fidelity on the Transfer-of-Training.



(a) Longitudinal cyclic control activity.

(b) Lateral cyclic control activity.



(c) Collective control activity.

(d) Pedals control activity.

Figure 7: Distribution of the average control activity for each group in each phase.

The results demonstrated the overall effectiveness of the training in both simulators, structured as a realistic flight lesson. Indeed, participants of both groups were able to stabilize the helicopter model, on average, in the 60% of the trials during the Evaluation phase. Moreover, no significant difference between CBT and CMS groups was found.

Although more experiment are needed to confirm the obtained results, the outcome of this experiment opens the possibility to replace or complement actual flight training hours with instruction hours on low-cost flight training devices. This can potentially reduce training costs and, eventually, pave the way towards a safety enhancement.

ACKNOWLEDGEMENTS

This study has been carried out in the context of the European Joint Doctorate NITROS (Network for Innovative Training on Rotorcraft Safety) project, whose main goal is to enhance rotorcraft safety by addressing critical aspects of their design. NITROS boasts the collaboration with important research partners. On behalf of the NITROS project, the authors would like to thank the Max Planck Institute for Biological Cybernetics for having supported the first author's secondment.
This project has received fundings from the European Union's Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement N° 721920.

REFERENCES

[1] David J. Allerton. The impact of flight simulation in aerospace. *The Aeronautical Journal*, 114 (1162):747–756, December 2010.

[2] Paul Adorian, W. N. Staynes, and Martin Bolton. The Evolution of the Flight Simulator. In *Proceedings of the Royal Aeronautical Society Conference, Fifty Years of Flight Simulation*, pages 1–23, Piccadilly Hotel, Piccadilly, London, UK, April 1979. Royal Aeronautical Society.

[3] L. D. Allen. Evolution of Flight Simulation. In *Proceedings of the AIAA Flight Simulation Technologies Conference, Monterey (CA)*, 1993.

[4] Ray L. Page. Brief History of Flight Simulation. In *Proceedings of the SimTecT 2000*, pages 1–11, 2000.

[5] Albert J. Rehmann, Robert D. Mitman, and Michael C. Reynolds. A Handbook of Flight Simulation Fidelity Requirements for Human Factors Research. Technical Report DOT/FAA/CT-TN95/46, U.S. Department of Transportation, Federal Aviation Administration Technical Center, Atlantic City International Airport, NJ 08405, December 1995.

[6] John E. Stewart II, William C. Barker, Dale S. Weiler, Jerry W. Bonham, and David M. Johnson. Assessing the Effectiveness of a Low-Cost Simulator for Instrument Training for the TH-67 Helicopter. Technical Report 1780, U.S. Army Research Institute for the Behavioral and Social Sciences, Fort Rucker, AL, December 2001.

[7] Henry L. Taylor, Gavan Lintern, Charles L. Hulin, Donald A. Talleur, Tom W. Emanuel Jr., and Sybil I. Phillips. Transfer of Training Effectiveness of a Personal Computer Aviation Training Device. *The International Journal of Aviation Psychology*, 9(4):319–335, 1999.

[8] Michael D. Proctor, Michael Panko, and Sharlene J. Donovan. Considerations for Training Team Situation Awareness and Task Performance Through PC-Gamer Simulated Multiship Helicopter Operations. *The International Journal of Aviation Psychology*, 14(2):191–205, 2004.

[9] Gustavo A. Ortiz. Effectiveness of PC-Based Flight Simulation. *The International Journal of Aviation Psychology*, 4(3):285–291, 1994.

[10] Michael D. Proctor, Maria Bauer, and Thomas Lucario. Helicopter Flight Training Through Serious Aviation Gaming. *The Journal of Defense Modeling and Simulation*, 4(3):277–294, July 2007.

[11] Davide Fabbroni, Stefano Geluardi, Carlo A. Gerboni, Mario Olivari, Giulia D'Intino, Lorenzo Pollini, and H. H. Bülthoff. Design of a Haptic Helicopter Trainer for Inexperienced Pilots. In *Proceedings of the AHS Annual Forum 73, Fort Worth (TX)*, pages 2097–2108, May 2017.

[12] Davide Fabbroni, Stefano Geluardi, Carlo A. Gerboni, Mario Olivari, Lorenzo Pollini, and H. H. Bülthoff. Quasi-Transfer-of-Training of Helicopter Trainer from Fixed-Base to Motion-Base Simulator. In *Proceedings of the 43rd European Rotorcraft Forum (ERF 2017)*, September 2017.

[13] Henry L. Taylor, Gavan Lintern, and Jefferson M. Koonce. Quasi-Transfer as a Predictor of Transfer from Simulator to Airplane. *The Journal of General Psychology*, 120(3):257–276, 1993.

[14] Frank M. Nieuwenhuizen and H. H. Bülthoff. The MPI CyberMotion Simulator: A Novel Research Platform to Investigate Human Control Behavior. *Journal of Computing Science and Engineering*, 7(2):122–131, June 2013.

[15] Stefano Geluardi. *Identification and augmentation of a civil light helicopter: transforming helicopters into Personal Aerial Vehicles*. MPI Series in Biological Cybernetics; 47, University of Pisa and Max Planck Institute for Biological Cybernetics, 2016.

[16] Stefano Geluardi, Frank M. Nieuwenhuizen, Joost Venrooij, Lorenzo Pollini, and H. H. Bülthoff. Frequency Domain System Identification of a Robinson R44 in Hover. *Journal of the American Helicopter Society*, 63(1):1–18, January 2018.

[17] Davide Fabbroni, Francesco Bufalo, Giulia D'Intino, Stefano Geluardi, Carlo A. Gerboni, Mario Olivari, Lorenzo Pollini, and H. H. Bülthoff. Transfer-of-Training: From Fixed- and Motion-base Simulators to a Light-Weight Helicopter. In *Proceedings of the AHS Annual Forum 74, Phoenix (AZ)*, May 2018.

[18] Lloyd D. Reid and Meyer A. Nahon. Flight Simulation Motion-Base Drive Algorithms. Part 1: Developing and Testing the Equations. Technical Report UTIAS 296, University of Toronto, Institute for Aerospace Studies, December 1985.

[19] Lloyd D. Reid and Meyer A. Nahon. Flight Simulation Motion-Base Drive Algorithms. Part 2: Selecting the System Parameters. Technical Report UTIAS 307, University of Toronto, Institute for Aerospace Studies, May 1986.

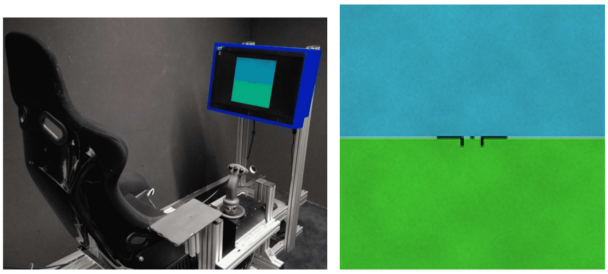
[20] *Unity User Manual*. Unity Technologies.

[21] Paul W. Caro. Aircraft Simulators and Pilot Training. *Human Factors*, 15(6):502–509, December 1973.

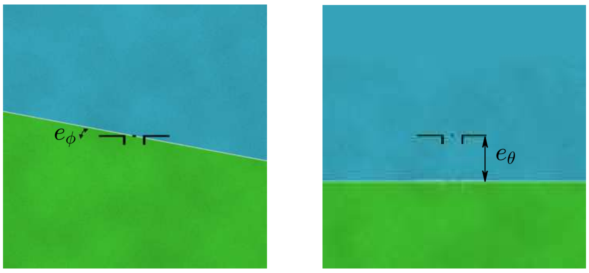
[22] Michael E. McCauley. Do Army Helicopter Training Simulators Need Motion Bases? Technical Report 1176, U.S. Army Research Institute for the Behavioral and Social Sciences, Fort Rucker, AL, February 2006.

A. RESULTS OF THE APTITUDE TEST

The aptitude test was performed in the Control Loading Lab (Fig. 8a) at the Max Planck Institute for Biological Cybernetics. A side-stick was used to give inputs to the controlled element. Roll and pitch axis of the side-stick were both active during the experiment. Therefore, both rotations (Fig. 8c) and translation (Fig. 8d) of the horizon marker on the artificial horizon were presented on the display. No other cues were presented.



(a) Apparatus. (b) Visual.



(c) Roll error. (d) Pitch error.

Figure 8: Experimental apparatus and visual used during the aptitude test.

The display used for the aptitude test is produced by VIEWPixx, VPixx Technologies Inc., Canada. The control device is an electrical control-loaded side-stick (Wittenstein Aerospace and Simulation GmbH, Germany). The sidestick was located on the right side of the chair where the participants were sitting. Thus, participants controlled the device using their right hand.

The aptitude test was composed of 10 trials. Each trial lasted 90 seconds. Thirty-three participants were tested. Their performances are presented in Fig. 9 as box-whiskers plots. From this figure, it can be noticed that starting from the 7th trial performances become stable as the median over the last 4 trials is almost constant and the between-subjects variability is smaller compared to the first trials.

A criterion, able to describe both overall and final behavior of each participant, was established to select participants. If over the last 4 trials, a participant had a number of trials within the 3rd quartile greater than 2, the participant was retained. If this number was lower than 2, the participant was excluded.

By applying this criterion, a total of five subjects were excluded. Two other participants were excluded because they did not comply with the safety requirements of the CyberMotion Simulator (CMS). Furthermore, two subjects left the study after the aptitude test due to personal reasons. The remaining twenty-four participants were ranked, based on their performance in the aptitude test, and methodically assigned to one of the two groups. From Fig. 10, it can be noticed that the two groups, on average, show equivalent performance for $\overline{\text{RMS}}$, $\overline{\text{RMS}}_\phi$ and $\overline{\text{RMS}}_\theta$. This is supported by independent-samples T tests for all three metrics (Tab. 6), which indicate that there is no significant between-group difference.

Table 6: Independent-samples T test to check difference between the two groups.

Metric	**t-test for Equality of Means**		
	t	df	Sig. (2-tailed)
$\overline{\text{RMS}}$	-0.039	22	0.969
$\overline{\text{RMS}}_\phi$	-0.371	22	0.714
$\overline{\text{RMS}}_\theta$	0.154	22	0.879

Hence, the metrics show that the groups have been equally distributed in terms of manual control skills throughout the aptitude test. The independent-samples T test was applied only after checking that data were approximately normally distributed and with homogeneous variance.

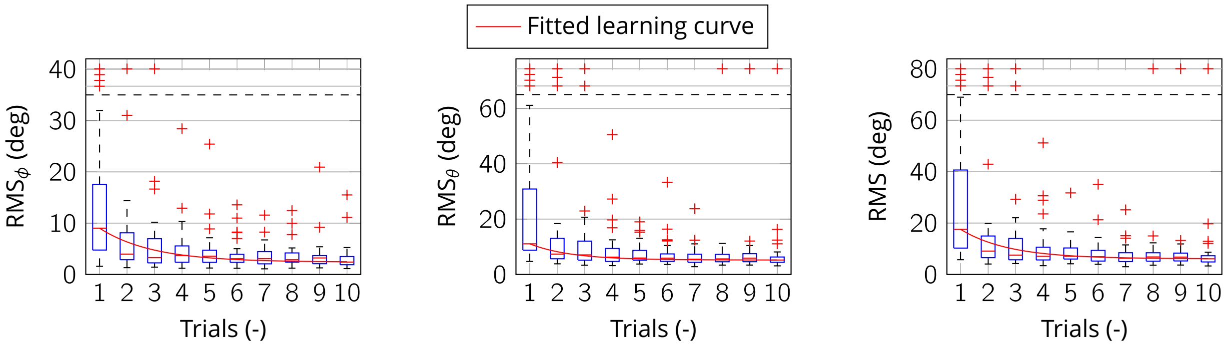
B. STUDENT HELPER

The part-task training was implemented in both simulators by using the software control system shown in Fig. 11. Here, the Helicopter Model to be controlled is a linear identified model of a Robinson R44 light-weight helicopter, described by the following state-space representation.

$$\dot{x} = \mathbf{A}x + \mathbf{B}\boldsymbol{u} \qquad (3)$$

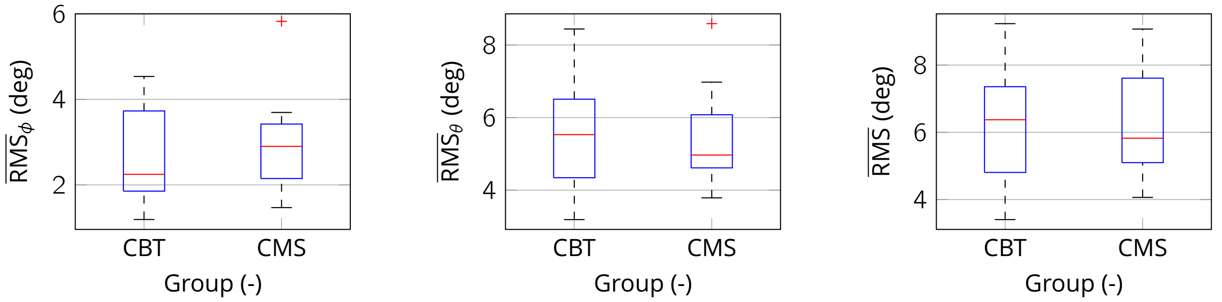
with $\boldsymbol{x} \in \mathbb{R}^{n_x=21}$, $\boldsymbol{u} = \begin{bmatrix} u_{lat} & u_{lon} & u_{ped} & u_{col} \end{bmatrix}^\mathsf{T} \in \mathbb{R}^{n_u=4}$.

The system of Eq. (3) is controlled by the combined action of student pilot $\boldsymbol{u_p}$ and software con-



(a) Roll RMS. (b) Pitch RMS. (c) Combined pitch and roll RMS.

Figure 9: Performances of the participants in the aptitude test.



(a) Roll RMS. (b) Pitch RMS. (c) Combined pitch and roll RMS.

Figure 10: Groups balance - Comparison of the average performances in the aptitude test.

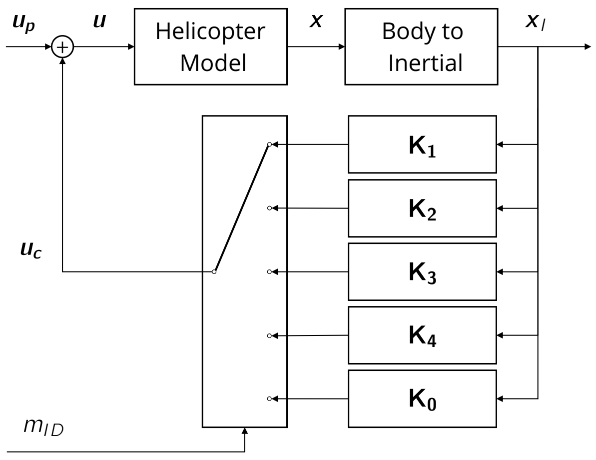


Figure 11: Logic of the controller used during the familiarization phase.

trol system $\boldsymbol{u_c}$ as in Eq. (4).

$$\boldsymbol{u} = \boldsymbol{u_p} + \boldsymbol{u_c} = \boldsymbol{u_p} - \mathbf{K}_i \boldsymbol{x_I} \tag{4}$$

The gain matrix $\mathbf{K}_i$ is the result of an optimization problem, based on the Linear Quadratic Regulator design implemented by Fabbroni et al.[11]. For each maneuver m_{ID}, a specific gain matrix was calculated:

$$\begin{aligned}
&\mathbf{K}_1 = \begin{bmatrix} \boldsymbol{k}_{1,1} \\ \boldsymbol{k}_{1,2} \\ 0 \\ \boldsymbol{k}_{1,4} \end{bmatrix} \quad \mathbf{K}_2 = \begin{bmatrix} \boldsymbol{k}_{2,1} \\ \boldsymbol{k}_{2,2} \\ \boldsymbol{k}_{2,3} \\ 0 \end{bmatrix} \quad \mathbf{K}_3 = \begin{bmatrix} \boldsymbol{k}_{3,1} \\ \boldsymbol{k}_{3,2} \\ 0 \\ 0 \end{bmatrix} \\
&\mathbf{K}_4 = \begin{bmatrix} 0 \\ 0 \\ \boldsymbol{k}_{4,3} \\ \boldsymbol{k}_{4,4} \end{bmatrix} \quad \mathbf{K}_0 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \quad \in \mathbb{R}^{n_u \times n_x}
\end{aligned} \tag{5}$$

with $\boldsymbol{k_{i,j}} \in \mathbb{R}^{1 \times n_x} \quad \forall i, j$.

Specifically, $\mathbf{K}_1$ is the gain matrix associated with the Hovering Turn maneuver, $\mathbf{K}_2$ with the Vertical Repositioning maneuver, $\mathbf{K}_3$ with the Vertical Repositioning, Heading Hold maneuver and $\mathbf{K}_4$ with the Hover A maneuver. Instead, $\mathbf{K}_0 = 0$ is associated with the Hover B maneuver, in which the student pilot is controlling the system with all control inputs.

Please note that in this setup the participants and the software control system never control the same channels at the same time.