This paper analyzes the effects of the helicopter dynamics on pilots’ learning process and transfer of learned skills during autorotation training. A quasi-transfer-of-training experiment was performed with fourteen experienced helicopter pilots in a moving-base flight simulator. Two types of helicopter dynamics, characterized by a different autorotative index, were considered: “hard”, with high pilot compensation required, and “easy”, with low compensation required. Two groups of pilots tested the two types of dynamics in a different training sequence: hard-easy-hard (HEH group) and easy-hard-easy (EHE group). Participants of both groups were able to attain adequate performance at touchdown in most of the landings with both types of dynamics. However, a clear positive transfer effect in terms of acquired skills is found in both groups from the hard to the easy dynamics, but not from the easy to the hard dynamics, confirming previous experimental evidence. Positive transfer is especially observed for the rate of descent at touchdown. The two groups differed in the control strategy applied, with the HEH group having developed a more robust control technique. During the last training phase the EHE group aligned its control strategy with that of the HEH group.

Nomenclature

\[
\begin{align*}
AI &= \text{Autorotation Index (ft}^3/\text{lb)} \\
c &= \text{Blade chord (ft or m)} \\
df &= \text{Number of degrees of freedom used in a statistical test}
\end{align*}
\]
\( F \) = \( F \)-statistic in the Analysis of Variance
\( h \) = Altitude (ft)
\( K \) = Scaling gain of the washout filter (-)
\( N_{des} \) = Number of landings within desired performance (- or % of the total number)
\( N_{ad} \) = Number of landings at least within adequate performance (- or % of the total number)
\( p \) = \( p \)-value (Significance)
\( q \) = Pitch rate (deg/s)
\( R \) = Main rotor radius (ft or m)
\( RMS \) = Root Mean Square
\( t \) = \( t \)-statistic in the t-test
\( t \) = Time (s)
\( V_x \) = Horizontal speed (kn)
\( V_z \) = Rate of descent (ft/min)
\( W \) = Helicopter weight (lb or kg)
\( U \) = \( U \)-statistic in the Mann-Whitney U test
\( Z \) = \( Z \)-statistic in the Wilcoxon signed-rank test
\( \delta_0 \) = Collective lever input (% of the available stroke)
\( \delta_{1s} \) = Longitudinal cyclic stick input (% of the available stroke)
\( \theta \) = Pitch angle (deg)
\( \zeta \) = Damping ratio of the washout filter (-)
\( \Omega \) = Main rotor speed (rad/s or % of the value at idle)
\( \omega_b \) = Third order break-frequency of the washout filter (rad/s)
\( \omega_n \) = Natural break-frequency of the washout filter (rad/s)

Subscripts
\( cush \) = Cushion
\( f \) = Failure
\( fl \) = Flare
\( reac \) = Reaction
\( rec \) = Recovery
\( rot \) = Rotation
\( td \) = Touchdown
\( x \) = Surge DOF of the washout filter
\( z \) = Heave DOF of the washout filter
\( \theta \) = Pitch DOF of the washout filter

Notation
\( \overline{\cdot} \) = Average value
\( \frac{d(\cdot)}{dt} \) = Time derivative
\( \text{sgn}(\cdot) \) = Sign function
\( \text{max}(\cdot) \) = Maximum value
\( |\cdot| \) = Absolute value

Acronyms

ADS = Aeronautical Design Standard
AI = Autorotation Index
ANOVA = Analysis of Variance
CED = Control Event Detection
CWA = Classical Washout Algorithm
DLP = Digital Light Processing
DOF = Degree of Freedom
E = Easy dynamics (low pilot compensation required)
EHE = Experiment group testing Easy-Hard-Easy dynamics
EHSAT = European Helicopter Safety Analysis Team
H = Hard dynamics (high pilot compensation required)
HEH = Experiment group testing Hard-Easy-Hard dynamics
HQR = Handling Qualities Rating
IHST = International Helicopter Safety Team
JHSAT = Joint Helicopter Safety Analysis Team
MAF = Moving Average Filter
MTE = Mission Task Element
qToT = quasi-Transfer-of-Training
RMS = Root Mean Square
RMSE = Root Mean Squared Error
RPM = Revolutions per minute
SIMONA = Simulation Motion Navigation
SRS  =  SIMONA Research Simulator
ToT  =  Transfer-of-Training

I. Introduction

Pilot training and flight simulators play a crucial role in rotorcraft safety. Accident analyses carried out by the U.S. Joint Helicopter Safety Analysis Team (JHSAT) [1, 2] and the European Helicopter Safety Analysis Team (EHSAT) [3, 4] pointed out that the development of a standardized training program for autorotation and emergency aircraft handling, as well as the improvement of simulator training for basic and advanced maneuvers, are essential. A combination of simulator and in-flight training is desirable, especially during the training of hazardous scenarios. The variety of conditions that pilots may face during emergencies requires experience and judgment in order to react promptly and avoid the many possible errors. Extensive practice is necessary to acquire robust and flexible piloting skills, giving the pilots the ability to adapt to different aircraft configurations and types. Because in-flight training is expensive and potentially dangerous, simulator training is the only viable alternative. Especially for rotorcraft, simulator usage has the potential to substantially reduce costs and risks, as shown by Memon et al. [5]. However, to avoid unrealistic training and negative transfer of skills when similar situations are encountered during actual flight, there is the need to bridge the gap between simulator scenarios and reality for edge-of-the-envelope flight.

Autorotation is considered to be a key critical training scenario [6–8]. It is a flight condition where the rotation of the rotor is sustained by the airflow moving up through the rotor, rather than by means of engine torque applied to the shaft. Helicopter pilots use autorotation following partial or total engine power failure to reach the nearest suitable landing site. The energy stored in the rotor is preserved at the expense of the helicopter’s potential energy, i.e., the altitude. Therefore, a helicopter can sustain autorotation only by means of descending flight.

One key challenge for obtaining representative simulation and effective pilot training in ground-based simulators is ensuring a sufficiently realistic flight model [9]. While this is true for flight simulation in general, the need for high-accuracy models is especially felt for flight conditions that are the result of abnormal modes of operation, such as autorotation. For example, the representation of the rotor wake plays an essential role in rotorcraft flight mechanics models. Houston and Brown [10] have investigated how vorticity transport models, as opposed to the simpler finite-state induced velocity models [11], affect modeling quality for autorotation. Furthermore, many studies [12–19] have focused on the effects of the variable rotor angular speed in autorotation, which are usually neglected in powered-flight rotorcraft models. Finally, key efforts were also devoted to model development and validation against wind tunnel and/or flight test data in autorotation and in the vortex ring state region for non-conventional helicopters, such as rotary decelerators of falling objects (e.g., ejection seat equipped with a folded rotor) [20] and coaxial helicopters [21, 22].

Besides model fidelity, pilots’ perception and their use of available cues is another aspect that should not be
underestimated in flight simulation for training purposes [23–25]. One of the most debated issues regarding simulator cueing is whether full-motion simulators are actually needed to achieve superior training quality. While this can be considered a general simulation question, there are peculiarities for each aircraft class, e.g. fixed-wing, helicopters, that require to perform a specific assessment for each one of them. Refs. [26] and [27] investigated the effects of a range of visual and motion cues settings on pilots’ ability to perform unpowered helicopter landings. Both works showed that helicopter touchdown performance, along with pilot opinion, improved with increased motion fidelity. However, in-simulator performance and simulator acceptance by the pilots are not metrics of training effectiveness when no transfer paradigm is adopted. To overcome the inconsistency among the results of the individual studies on the need of motion bases [28], de Winter et al. [29] conducted a meta-analysis on 24 transfer-of-training experiments with motion as an independent variable, using both fixed- and rotary-wing aircraft models, showing that there is no overall evidence that motion improves performance in real aircraft, even though positive effects in favor of motion are observed in quasi-transfer studies and for individuals without flight experience learning disturbance-rejection tasks or maneuvers of vehicles with low dynamic stability, such as helicopters. Since then, other transfer-of-training experiments were performed for both fixed-wing aircraft [24, 30] and helicopters [31–34]. Most of these experiments use a quasi-transfer paradigm, where trained skills are applied on a simulator with capabilities that are beyond those of a typical training simulator [35], corroborating the assumption that such simulators act as a valid replacement for the actual aircraft. Similar results are obtained in all these studies: the need of simulators with a motion system cannot be claimed. On the contrary, subjects who trained in poorer cueing situations developed control strategies that were revealed to be adaptable to higher fidelity conditions.

Despite these recent efforts devoted to achieving more accurate rotorcraft models and clarifying the relation between simulator cueing and training effectiveness, only a few studies have explicitly investigated the effects of rotorcraft model fidelity and dynamics variations on pilot behaviour and (transfer of) training, e.g., [36–38]. Especially for a critical hands-on maneuver such as autorotation, pilots need to adjust their control strategy according to the helicopter dynamics they control [35, 39, 40]. Helicopters with different handling characteristics may require very different skills from pilots to accomplish the task. Earlier experiments in training a lateral sidestep hover maneuver [36], showed that flight-naïve participants – i.e., without any previous real or simulated flight experience – are more likely to develop robust and flexible flying skills when they start the training in a helicopter with agile system dynamics. According to Nusseck et al. [36], starting the training of a certain task with the most challenging configuration provides the pilot with the ability to accomplish the same task with every other configuration after a short adaptation phase.

This result is consistent with the principle of perceptual learning, which Gold et al. [41, (p. R46)] define as “experience-dependent enhancement of our ability to make sense of what we see, hear, feel, taste or smell. These changes are permanent or semi-permanent, as distinct from shorter-term mechanisms like sensory adaptation or habituation. Moreover, these changes are not merely incidental but rather adaptive and therefore confer benefits, like improved
sensitivity to weak or ambiguous stimuli”. Thus, training with the highest resource-demanding setting enhances perceptual learning, and the improved perception skills allow pilots trained in this setting to more easily adjust their control strategy when transferred to another, easier condition.

This paper investigates whether the acquisition of flying skills for autorotation, and their transfer, are affected by the helicopter dynamics. We hypothesized that certain dynamics may lead to the development of a more robust control behavior, one that can be easily adapted to a helicopter with different dynamics, yielding substantial benefits in terms of engine failure handling capabilities.

The results of a quasi-Transfer-of-Training (qToT) experiment with fourteen experienced helicopter pilots, divided in two groups, performed in TU Delft’s SIMONA Research Simulator (SRS) are presented to corroborate this hypothesis. Several metrics have been used to compare the performance at touchdown. Additionally, a novel method, referred to as Control Event Detection (CED), is presented, to allow an in-depth analysis of pilot control actions involved in a successful autorotative landing.

The paper is structured as follows. Section II describes the experimental design and set-up. Results are presented in Section III and discussed in Section IV. Conclusions are drawn in Section V.

II. Methods

A. Task

In rotorcraft handling quality research, experimental tasks are usually defined according to the specifications of the mission-oriented design standard, the ADS-33E [42]. Although conceived for military rotorcraft, the ADS-33E are widely used to assess handling qualities characteristics of commercial rotorcraft as well, as there is no counterpart in the civil domain. However, the use of ADS-33E Mission Task Elements (MTEs) is not always relevant, especially in the design of training tasks. Furthermore, the ADS-33E does not have a specific Autorotation Maneuver MTE. Since there are no specific handling quality metrics for autorotation, pilot-in-the-loop autorotation maneuvers are usually evaluated based on subjective pilot feedback and comments and on objective measurements of landing survivability metrics [43].

For this experiment, a MTE was defined for the straight-in autorotation maneuver; the proposed test course is shown in Fig. 1. The simulation starts with the helicopter trimmed in straight level flight at 60 knots air speed, at an altitude of 1,000 ft. The symmetry plane of the helicopter is aligned with the center line of a runway, whose starting point is located 3,281 ft ahead the helicopter initial position. The pilot has to keep constant speed and altitude until the power failure is triggered from the control room. As soon as the pilot recognizes the unannounced failure, he has to recover starting a steady descent in autorotation, maintaining 60 knots air speed and keeping the rotor RPM in the green arc of the tachometer. When close enough to the ground the pilot has to flare, to reduce both the rate of descent and the forward speed and finally level the skids with the ground, to avoid tail strike, and pull-up the collective to cushion the
touchdown. The contact accelerations at touchdown were not modeled. Therefore, the simulation stopped automatically once the center of gravity of the helicopter reached 6.6 ft above the ground.

![Diagram showing the suggested course for straight-in autorotation maneuver.](image)

(a) Side view.

(b) Top view.

**Fig. 1 Suggested course for straight-in autorotation maneuver.**

Performance standards for the straight-in autorotation maneuver are adapted from Sunberg et al. [43, 44] and are listed in Tab. 1. The values of the horizontal speed and of the rate of descent at touchdown refer to the AH-1G helicopter [43], which has a similar skid landing gear as the baseline helicopter (Bo-105) considered in this paper. Therefore, these were not changed. Although characterized by a similar landing system, the AH-1G and the Bo-105 are different helicopters, with different performance and intended role. Indeed, the AH-1G is a two-blades rotor, single-engine attack helicopter, whereas the Bo-105 is a light, twin-engine, multi-purpose helicopter with a four-blades hingeless rotor. The maximum values of the pitch angle at touchdown, which are responsible of preventing tail strike, were slightly increased due to the different helicopter geometry. Desired performance translates into a successful landing, i.e., the helicopter’s final state at ground contact is such that the aircraft and crew survivability are not threatened. Adequate performance translates into marginal landing conditions, that would likely result in damage to the aircraft, but be survivable to the occupants and the equipment. The values presented in Tab. 1 are defined according to landing survivability metrics that are based on specifications for military helicopters’ structural design [45, 46] and on the accident analysis conducted by Crist and Symes [47].
### Table 1  Performance – Straight-in Autorotation Maneuver (adapted from Sunberg et al. [43]).

<table>
<thead>
<tr>
<th>Metric</th>
<th>Performance</th>
<th>Desired</th>
<th>Adequate</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Minimum</td>
<td>Maximum</td>
<td>Minimum</td>
</tr>
<tr>
<td>Pitch angle at touchdown $\theta_{td}$ (deg)</td>
<td>$-5$</td>
<td>$12$</td>
<td>$-5$</td>
</tr>
<tr>
<td>Horizontal speed at touchdown $V_{xt_d}$ (kn)</td>
<td>$0$</td>
<td>$30$</td>
<td>$0$</td>
</tr>
<tr>
<td>Rate of descent at touchdown $V_{z_{td}}$ (ft/min)</td>
<td>$0$</td>
<td>$480$</td>
<td>$0$</td>
</tr>
<tr>
<td>Pitch rate at touchdown $q_{td}$ (deg/s)</td>
<td>$-30$</td>
<td>$20$</td>
<td>$-50$</td>
</tr>
</tbody>
</table>

### B. Helicopter Dynamics

Participants performed the straight-in autorotation task by controlling a four degrees-of-freedom (3-DOF longitudinal dynamics plus rotorspeed DOF), non-linear and generic helicopter model with quasi-steady flapping dynamics [19]. This generic model can be used in combination with different parameters sets to approximate the dynamic response of any conventional helicopter configuration. Since the final part of the autorotation is mainly a longitudinal maneuver, the use of a 3-DOF symmetrical helicopter model is acceptable.

Scaramuzzino et al. [19] compared the stability characteristics in autorotation of helicopters with a different autorotative flare index (AI) [48]. The AI can be interpreted as the ratio between the available energy, i.e., rotor kinetic energy $I_R \Omega^2 / 2$, where $I_R$ is the polar moment of inertia of the rotor system and $\Omega$ is the rotor RPM, and the energy required to stop the rate of descent of the helicopter, proportional to the helicopter weight $W$ and the disk loading $DL$. Therefore, high values of the index are desirable. In order to compare the values of this index for various helicopters, it is convenient to plot the parameter proportional to rotor kinetic energy per unit gross weight $I_R \Omega^2 / 2 / W$ versus disk loading $DL$. This graphical form is adopted in Fig. 2, where an overview of typical values of the autorotation index is given. Several helicopters have been considered and all of them have an autorotative index between 5 and 40 ft$^3$/lb. However, it can be noted that values of the index above 30 ft$^3$/lb can only be achieved by single-engine helicopters, whereas values below 15 ft$^3$/lb are typical of large helicopters (maximum mass greater than 9072 kg).

From the wide range of configurations studied by Scaramuzzino et al. [19], two were selected for the current experiment. The “hard” dynamics is representative of the Bo-105 helicopter and was taken from Padfield [50]. The “easy” dynamics represents a variation of the Bo-105 helicopter with reduced weight in order to achieve a higher AI.

Fig. 3 shows a comparison of the dynamics modes in steady descent in autorotation for the two helicopter configurations considered here. The rate of descent for the hard configuration is 1630.2 ft/min, whereas for the easy configuration is 1649.8 ft/min. Although the modes of the two configurations do not differ considerably, in particular in terms of the pitch subsidence, it can be noted that the phugoid mode of the hard dynamics is less stable than that of the easy dynamics, whereas the opposite holds for the rotorspeed mode. While similar in terms of stability characteristics,
Fig. 2  Autorotative indices at standard sea level conditions for several helicopters (revised from Fradenburgh [48] and Leishman [49]).

these two configurations proved to be considerably different in terms of handling qualities during a pre-experiment with a test pilot, both concerning objective metrics of performance at touchdown (Tab. 1) and subjective handling quality ratings provided by the pilot. The selection process of these two helicopter dynamics is extensively discussed in Appendix A.

Fig. 3  Comparison between modes in steady descent in autorotation for the two different helicopter’s configurations at standard sea level conditions.

C. Experiment Structure

The experiment is structured as in Tab. 2 and consists of four phases:
1) Familiarization: this phase was intended to help the participants get acquainted with the simulation environment (helicopter model, cockpit ergonomy, control inceptors, etc.). For this reason, the simulator motion system was disabled and each participant performed the task with either the hard or the easy helicopter dynamics. These runs were not used in the analysis.

2) Training: each participant performed the task with the same helicopter dynamics used during the Familiarization phase. Starting from this session, the simulator motion system was enabled.

3) Transfer: each participant performed the task with the other helicopter configuration.

4) Back-transfer: each participant performed the task with the initial hard/easy helicopter configuration.

In total, the whole experiment session for each participant lasted approximately 3 hours.

Table 2  Experiment phases.

<table>
<thead>
<tr>
<th>Phase</th>
<th>HEH group</th>
<th>EHE group</th>
<th>Duration (number of autorotative landings)</th>
<th>Motion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Familiarization</td>
<td>Hard helicopter dynamics</td>
<td>Easy helicopter dynamics</td>
<td>3</td>
<td>Off</td>
</tr>
<tr>
<td>Training</td>
<td>Hard helicopter dynamics</td>
<td>Easy helicopter dynamics</td>
<td>15</td>
<td>On</td>
</tr>
<tr>
<td>Transfer</td>
<td>Easy helicopter dynamics</td>
<td>Hard helicopter dynamics</td>
<td>15</td>
<td>On</td>
</tr>
<tr>
<td>Back-Transfer</td>
<td>Hard helicopter dynamics</td>
<td>Easy helicopter dynamics</td>
<td>14*</td>
<td>On</td>
</tr>
</tbody>
</table>

* The log file of the 15th run of the back-transfer phase of some participants was corrupted. Due to the loss of these data, the 14th run was considered as the last run of the back-transfer phase.

D. Dependent measures

To investigate the effect of the helicopter dynamics (independent variable) on autorotation performance and training, the dependent measures related to the MTE definition presented in Tab. 1 were considered. Since those measures assess only the performance at touchdown, other metrics were also taken into account to compare the control strategies adopted by the participants of the two experiment groups, namely:

1) Number of landings at least within adequate performance (Tab. 1).
2) Number of landings within desired performance (Tab. 1).
3) Reaction time: time required by the pilot to lower the collective after engine failure (Fig. 1).
4) Recovery time: time required by the pilot to enter a steady descent in autorotation (Fig. 1).
5) Flare initiation altitude: altitude at which the pilot initiates the flare by pulling back the cyclic stick (Fig. 1).
6) Rotation altitude: altitude at which the pilot levels the skids with the ground by pushing forward the cyclic stick (Fig. 1).
7) Cushion altitude: altitude at which the pilot raise the collective to cushion the touchdown (Fig. 1).
Metrics 3 to 7 were extracted from the experiment time histories using a newly developed methodology called Control Event Detection (CED), which is presented in Section II.I.1.

E. Hypotheses

For this experiment only one main hypothesis was tested. Based on previous experimental evidence [36] and on current in-flight training procedures, it is envisioned that pilots who start the training with the most challenging configuration (hard dynamics), are more likely to develop robust and flexible autorotation skills that can be easily adapted to different helicopter configurations and dynamics. Therefore, it is expected that flying skills are positively transferred from the hard to the easy dynamics, but not conversely. When positive transfer happens, we expect to see lower rates of descent after transition to a different dynamics, as a lower descent rate is a key indicator for a controlled and smooth touchdown [8]. Among all the dependent measures, the rate of descent is thus expected to cover a key role to corroborate our hypothesis.

F. Participants

A total of fourteen experienced helicopter pilots with a different background (license type), with a mix of civil and military experience and with a different in-flight and simulator experience took part in the experiment; all of them were male. The participants had an average age of 41.71 years ($\sigma = \pm 9.50$ years) and an average helicopter experience of 2,669 flight hours ($\sigma = \pm 2,336$ flight hours), ranging from a minimum of 120 to a maximum of 6,100 flight hours. Participants were divided in two groups in such a way that they had, on average, a comparable number of flight hours and a similar distribution, as shown in Tab. 3. Beside the number of flight hours, also pilots background was considered during the separation of the pilots in the two groups.

<table>
<thead>
<tr>
<th>Participant ID</th>
<th>HEH group</th>
<th>EHE group</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Age</td>
<td>Flight Hours</td>
</tr>
<tr>
<td>1</td>
<td>40</td>
<td>3000</td>
</tr>
<tr>
<td>2</td>
<td>43</td>
<td>6100</td>
</tr>
<tr>
<td>3</td>
<td>32</td>
<td>120</td>
</tr>
<tr>
<td>4</td>
<td>40</td>
<td>1500</td>
</tr>
<tr>
<td>5</td>
<td>28</td>
<td>1900</td>
</tr>
<tr>
<td>6</td>
<td>35</td>
<td>2000</td>
</tr>
<tr>
<td>7</td>
<td>38</td>
<td>7000</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>avg</th>
<th>std</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>HEH group</td>
<td>36.6</td>
<td>5.2</td>
<td>46.9</td>
</tr>
<tr>
<td>EHE group</td>
<td>3089</td>
<td>2527</td>
<td>2249</td>
</tr>
</tbody>
</table>

Participants signed an informed consent prior to the experiment. The experiment has been approved by the Human Research Ethics Committee of Delft University of Technology under the approval letter number 940.
G. Apparatus

The experiment was conducted in the SIMONA Research Simulator (SRS) (Fig. 4), which is a moving-base simulator at the Faculty of Aerospace Engineering of TU Delft [51]. The SRS is equipped with a 6 degrees-of-freedom hydraulic motion system, which was used in the experiment to provide motion cues.

![The SIMONA Research Simulator at Delft University of Technology.](image)

In terms of visual equipment, the SRS is fitted with a 180° × 40° 3-projector Digital Light Processing (DLP®) collimated display. A representative out-of-the-window scenery was presented on this display (Fig. 5a). Furthermore, an instrument panel (Fig. 5b), consisting of a tachometer, airspeed indicator, artificial horizon, altimeter, yaw string, compass and vertical speed indicator, and a trim display (Fig. 6) were projected on two monitors inside the cockpit. Pilots used the trim display only before the start of each run in order to find the trim position of all the flight controls. This enables them to keep the initial equilibrium condition (straight level flight at 60 kn) and avoid a transient response to reestablish it.

![Out-of-the-window scenery and instrument panel used for the current experiment.](image)

The right seat of the cockpit was equipped with a realistic helicopter control inceptor with programmable control loading system, whose parameters were set as reported in Tab. 4 after consultation with test pilots [52]. Since the
The helicopter model used in the experiment [19] only consists of the longitudinal dynamics, solely the longitudinal cyclic stick and the collective lever were to be controlled by the participants. Therefore, the lateral cyclic stick and the rudder pedals were not used. Lateral cyclic settings are also presented in Tab. 4 because the cyclic stick was not constrained to move only longitudinally. Pilots could also move the cyclic stick laterally, but this would not have produced any effect on the flight dynamics model response.

### Table 4 Control loading settings.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Longitudinal Cyclic</th>
<th>Lateral Cyclic</th>
<th>Collective</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Periodic</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Forward friction level (N)</td>
<td>2.0</td>
<td>2.0</td>
<td>6.0</td>
</tr>
<tr>
<td>Positive forward stop (deg)</td>
<td>15.0</td>
<td>15.0</td>
<td>16.0</td>
</tr>
<tr>
<td>Negative forward stop (deg)</td>
<td>−15.0</td>
<td>−15.0</td>
<td>−16.0</td>
</tr>
<tr>
<td><strong>Non periodic</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Linkage stiffness (N/deg)</td>
<td>50.0</td>
<td>50.0</td>
<td>50.0</td>
</tr>
<tr>
<td>Linkage damping (N s/deg)</td>
<td>0.01</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>Positive aft travel limit (deg)</td>
<td>14.8</td>
<td>14.8</td>
<td>15.8</td>
</tr>
<tr>
<td>Negative aft travel limit (deg)</td>
<td>−14.8</td>
<td>−14.8</td>
<td>−15.8</td>
</tr>
<tr>
<td>Aft friction (N)</td>
<td>2.0</td>
<td>2.0</td>
<td>6.0</td>
</tr>
<tr>
<td>Aft inverse damping (deg/N/s)</td>
<td>10.0</td>
<td>10.0</td>
<td>10.0</td>
</tr>
<tr>
<td>Second feel spring slope (N/deg)</td>
<td>3.0</td>
<td>3.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Breakout level (N)</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
</tbody>
</table>

Rotor sound was played during the simulation to increase immersion. The sound was modulated based on the value of the rotor RPM, so that the participant could use sound cues as a source of information to control the rotor RPM, rather than by looking at the instrument panel. Moreover, a low-RPM acoustic warning was activated every time the rotorspeed dropped below 85%. The low-RPM warning was used as a backup cue for the rotor sound, so that the failure could be recognized without necessarily looking at the instruments. Engine sound was not included.
H. Motion Filter Tuning

The Classical Washout Algorithm (CWA) is used to map the vehicle motion on the simulator workspace [53]. In the experiment, only three degrees of freedom (longitudinal dynamics) are active. Therefore, the CWA reduces to a set of three high-pass filters in the pitch, surge and heave axes, respectively. These filters were selected to be of second order for the pitch and surge axes, and of third order for the heave axis. Although surge and heave axes are both translational degrees of freedom, a different order of the filter was selected for these two axes. Indeed, a second order high-pass filter along the surge axis allows to achieve sufficient washout through the use of tilt coordination. This was first observed by Reid and Nahon [54] and reiterated by Grant [55]. Therefore, the combination of the tilt coordination and the body to inertial transformation effectively adds one order of washout. Each second order filter has three adjustable parameters, namely, the scaling gain $K$, the damping ratio $\zeta$, and the natural break-frequency $\omega_n$. The third order filter features an additional parameter, which is the third order break-frequency $\omega_b$.

The scaling gain $K$ and the damping ratio $\zeta$ were fixed for all simulator degrees of freedom at a value of 0.5 and 0.7071, respectively. The third order break-frequency $\omega_b$ of the heave axis filter was fixed at a value of 0.2 rad/s. These values were chosen in such a way that the respective degree of freedom does not require a large amount of the motion space, but provides at the same time cues of sufficient fidelity to the pilot to perform the training task. Indeed, a homogeneous scaling gain of 0.5 is the maximum value that can be attained for this maneuver. A two-pole filter with a damping ratio of 0.7071 is a second-order Butterworth filter, which assures a maximally flat magnitude response (sharpest roll-off possible without inducing peaking at the pole frequency) and, at the same time, a relatively smooth step response in the time domain (limited overshoot). Furthermore, the natural break-frequencies of the pitch and surge axes were constrained to be equal. These assumptions reduce the number of tuning parameters to two: the natural break-frequencies of the pitch and heave axes.

The optimal motion settings for the natural frequencies were determined using a two-dimensional grid search, based on the analysis proposed by Gouverneur et al. [56]. The final motion filter settings for the 3 degrees-of-freedom are presented in Tab. 5 and are based on ten data sets from a pilot-in-the-loop pre-experiment. These data sets are representative of very aggressive pilot control behavior and were used to perform a conservative motion tuning, so that the simulator would stay within its physical limits for most of the pilots. Fig. 7 shows the critical combinations of heave and surge frequencies for each of the ten data-sets collected during the pre-experiment, indicating the limit bounds below which either a motion system actuator length limit ($\pm 0.575$ m), speed limit ($\pm 1$ m/s) or both were exceeded. The selected setting is located above all the limit bounds.
Table 5  Motion cueing settings.

<table>
<thead>
<tr>
<th>DOF</th>
<th>$K$ (-)</th>
<th>$\omega_n$ (rad/s)</th>
<th>$\zeta$ (-)</th>
<th>$\omega_n$ (rad/s)</th>
<th>Order (-)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Heave</td>
<td>0.5</td>
<td>3.5</td>
<td>0.7071</td>
<td>0.2</td>
<td>3</td>
</tr>
<tr>
<td>Surge</td>
<td>0.5</td>
<td>1.5</td>
<td>0.7071</td>
<td>0.0</td>
<td>2</td>
</tr>
<tr>
<td>Pitch</td>
<td>0.5</td>
<td>1.5</td>
<td>0.7071</td>
<td>0.0</td>
<td>2</td>
</tr>
</tbody>
</table>

Grid  
- Critical bounds run 1  
- Critical bounds run 2  
- Critical bounds run 3  
- Critical bounds run 4  
- Critical bounds run 5  
- Critical bounds run 6  
- Critical bounds run 7  
- Critical bounds run 8  
- Critical bounds run 9  
- Critical bounds run 10  
- Selected setting

![Fig. 7 Two-dimensional grid search, based on the analysis proposed by Gouverneur et al. [56] of ten different runs. The black dot indicates the motion filter setting used in the experiment.](image_url)

I. Data Processing and Analysis

1. Control Event Detection Methodology

The transition between two phases of a maneuver is characterized by a change in the pilot’s control strategy. In order to determine when this change actually happens, a pilot’s control inputs need to be filtered to remove the high frequency deflections related to the small inputs that are applied by the pilot to correct for small perturbations and deviations from the desired flight condition. According to the authors’ knowledge, the Moving Average Filter (MAF) is the only available technique to find the exact transition time between two consecutive phases of a maneuver. This technique has been used to determine the flare initiation time during landings with fixed-wing aircraft [57] and is grounded on the fact that the aircraft phugoid eigenmotion is excited when the pilot initiates the flare by gradually pulling back the control column to increase the elevator deflection. Since this technique is tailored for fixed-wing aircraft, it often fails when applied to rotorcraft. Therefore, a more general method, called Control Event Detection (CED) Methodology, has been developed and proposed here for the first time. This method is applied in this paper to identify different phases of the autorotation maneuver (Fig. 1), but is envisioned to be applicable to other maneuvers and/or to different vehicles/system dynamics as well. For our current analysis, the outcomes of the CED methodology were verified using visual inspection.
of our data: information deriving from time histories of both control inputs and helicopter states was combined to assess the correctness of the points identified by the CED methodology (e.g., for the flare initiation altitude we focus our attention on the last 150/200 ft altitude of the time history and we look at the longitudinal cyclic stick input, the rotor angular speed, the rate of descent, and the forward speed).

This method only requires that the input is positive. This can be easily achieved by defining the current value of the input as a percentage of the available stroke. The procedure is composed of the following steps:

1) Computation of the time derivative of the control input (Fig. 8b).

2) Identification of the sign of the time derivative of the control input (Fig. 8c).

3) Intervals in which the sign of the time derivative of the control input is zero and that are bounded both from the left and the right by positive (negative) intervals, are turned into positive (negative) intervals (Fig. 8d). This step is required only for noisy control inputs.

4) Multiplication of the control input by the sign of its time derivative (Fig. 8e). In this way, the control input is divided into intervals, where it either assumes strictly monotone values or is zero.

5) For each strictly monotone interval, the Root Mean Squared Error (RMSE) of the control input with respect to the initial value of that interval is calculated. In this way, it is possible to quantify the control activity during each interval, by considering the deviation of the control with respect to the initial position.

6) Multiplication of the RMSE of each interval by the sign of the time derivative of the control input (Fig. 8e).

The different steps of the method are shown graphically in Fig. 8, where the time history after the failure ($t_f$ is the time when the failure happens) of the cyclic stick is analyzed for one experiment run.

Visual inspection can be used to identify the instants of interest, that are characterized by higher RMSE with respect to neighboring regions. In case a large number of time series has to be processed, visual inspection can be used for a sample of them to train a sequence-to-sequence deep neural network. The trained network can be subsequently applied to recognize the instants of interest for the remainder time series. As an alternative to deep learning, the instants of interest can be identified by isolating specific regions and defining threshold for the RMSE.

The CED method was applied to all the experiment runs in order to extract information about pilots’ control strategy and activity from the time histories. For example, the reaction time and the cushion altitude are identified by applying the CED Methodology to the collective lever input. The flare initiation altitude and the rotation altitude are identified by analyzing with the same method the cyclic stick input. The identification of the recovery time is based on the analysis of both the collective lever and cyclic stick inputs.

2. Statistical Analysis

Prior to performing the statistical tests, all the dependent measures defined in Section II.D, except the number of landings at least within adequate performance and that within desired performance, were averaged over the last 10 runs.
of each phase for every participant. Mixed repeated measures ANOVA (Analysis of Variance) tests were conducted on all the dependent variables, considering the experiment phases as main within-subjects factor, characterized by three levels: training, transfer and back-transfer, and the groups as main between-subjects factor, characterized by two levels: HEH and EHE. Before conducting the statistical tests, the fulfillment of the ANOVA assumptions, i.e., normality and sphericity, was verified. Regarding the normality assumption, some skewness in the data was accepted, provided that it was nothing too alarming. Indeed, the ANOVA is a robust technique and should still provide reliable results also in presence of minor violations. For variables in which sphericity was violated according to Mauchly’s test, we adopted either the Greenhouse-Geisser correction, when the Greenhouse-Geisser estimate of sphericity $\varepsilon$ is below 0.75 ($\varepsilon < 0.75$), or the Huynd-Feldt correction, when the Greenhouse-Geisser estimate of sphericity $\varepsilon$ is above 0.75 ($\varepsilon > 0.75$).

In the event of a statistically significant interaction effect between the two main factors (within- and between-subjects factors), main effects may provide misleading information [58]. Therefore, the so-called simple main effects are investigated. Dependent-samples t-tests between the phases of the experiment for each group are used to investigate the simple within-subjects effect and independent-samples t-tests between the groups in each phase are adopted to examine the simple between-subjects effect. If the interaction effect between the two main factors is not statistically significant, the two main effects are analyzed and if either of them is statistically significant, the respective simple main effect is investigated accordingly.
Fig. 8  Graphical summary of the Control Event Detection Methodology applied to the cyclic stick input.
III. Results

Results are presented in the following figures as box-whiskers plots. On each box, the horizontal line 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 represented by cross markers.

Statistically significant results of the t-tests are shown as follows. A curly brace with an asterisk on top is used to indicate a statistically significant difference between the two groups in a specific phase of the experiment. A curved arrow with an asterisk on the left indicates a statistically significant transfer of training for a specific group.

A. Performance Scores

Tab. 6 summarizes the results of a repeated measures ANOVA test for the different dependent measures considered in this study.

Table 6  Repeated Measures ANOVA results for all the dependent variables.

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>Between-subjects Factor (Group)</th>
<th>Within-subjects Factor (Phase)</th>
<th>Interaction (Phase*Group)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nadm</td>
<td>df 1 1.155 F 0.304 Sig.</td>
<td>df 2 3.043 F 0.066 Sig.</td>
<td>df 2 0.406 F 0.671 Sig.</td>
</tr>
<tr>
<td>Ndes</td>
<td>df 1 0.043 F 0.838 Sig.</td>
<td>df 2 3.796 F 0.037* Sig.</td>
<td>df 2 2.502 F 0.103 Sig.</td>
</tr>
<tr>
<td>θstd</td>
<td>df 1 0.394 F 0.542 Sig.</td>
<td>df 2 1.270 F 0.299 Sig.</td>
<td>df 2 0.098 F 0.907 Sig.</td>
</tr>
<tr>
<td>Vstd</td>
<td>df 1 0.001 F 0.971 Sig.</td>
<td>df 2 2.040 F 0.152 Sig.</td>
<td>df 2 1.103 F 0.348 Sig.</td>
</tr>
<tr>
<td>Vzstd</td>
<td>df 1 0.397 F 0.541 Sig.</td>
<td>df 2 4.789 F 0.018* Sig.</td>
<td>df 2 1.472 F 0.249 Sig.</td>
</tr>
<tr>
<td>qtd</td>
<td>df 1 6.237 F 0.028* Sig.</td>
<td>df 2 0.461 F 0.636 Sig.</td>
<td>df 2 0.863 F 0.434 Sig.</td>
</tr>
<tr>
<td>Ωt d</td>
<td>df 1 3.198 F 0.099 Sig.</td>
<td>df 2 0.249 F 0.782 Sig.</td>
<td>df 2 1.926 F 0.168 Sig.</td>
</tr>
<tr>
<td>Ωt d</td>
<td>df 1 9.753 F 0.009* Sig.</td>
<td>df 2 1.818 F 0.184 Sig.</td>
<td>df 2 0.089 F 0.915 Sig.</td>
</tr>
<tr>
<td>Ωt rec</td>
<td>df 1 0.165 F 0.692 Sig.</td>
<td>df 2 0.502 F 0.611 Sig.</td>
<td>df 2 0.501 F 0.612 Sig.</td>
</tr>
<tr>
<td>Ωt rec</td>
<td>df 1 1.466 F 0.249 Sig.</td>
<td>df 2 0.188 F 0.830 Sig.</td>
<td>df 2 0.052 F 0.949 Sig.</td>
</tr>
<tr>
<td>Ωt cash</td>
<td>df 1 5.158 F 0.042* Sig.</td>
<td>df 2 0.673 F 0.520 Sig.</td>
<td>df 2 1.123 F 0.342 Sig.</td>
</tr>
<tr>
<td>Ωt id</td>
<td>df 1 2.841 F 0.118 Sig.</td>
<td>df 2 1.702 F 0.204 Sig.</td>
<td>df 2 1.244 F 0.306 Sig.</td>
</tr>
</tbody>
</table>

* Statistically significant (p ≤ 0.05) difference between compared samples.

Fig. 9 shows the number of landings at least within adequate performance and those within desired performance for each group in each phase. Fig. 9a illustrates that participants of both groups were able to attain at least adequate performance, i.e., a survivable landing, in most of the experiment runs. However, the EHE group shows higher within group variability than the HEH group. This is particularly true for the training and the back-transfer phases, where the EHE group controls the easy dynamics.

Participants of both groups struggled to attain desired performance (i.e., a successful landing), as shown in Fig. 9b. Nonetheless, the success rate for the participants of the HEH group increases during the transfer phase and decreases
again during the back-transfer phase to values comparable to those of the training phase. Therefore, participants of the HEH group are able to more easily attain desirable performance with the easy configuration.

In contrast, there is no substantial difference in the number of successful landings between the training and the transfer phase for the participants of the EHE group, whereas there is an increase in the success rate during the back-transfer phase. Fig. 9b suggests that for the HEH group there is positive transfer of training from the hard configuration to the easy configuration and no transfer of training from the easy configuration to the hard configuration. Although not statistically significant, a similar trend is observed also for the EHE group.

![Fig. 9](image_url)

**Fig. 9** Distribution of the number of landings at least within adequate performance and those within desired performance for each group in each phase.

Tab. 6 highlights a statistically significant difference in terms of number of successful landings between the experiment phases (within-subjects effect: $F(2,24) = 3.796, p = 0.037$). The presence of an overall significant effect in the complete data set of the number of successful landings was further investigated by performing t-tests on individual sets of samples. Tab. 7 illustrates the results of these tests. A significant difference in terms of number of successful landings only occurs between the training and the transfer phase for the HEH group (from the hard to the easy helicopter dynamics: $t(6) = −3.813, p = 0.009$). This partially confirms what has already been observed from Fig. 9b.

The effectiveness of the training was further investigated by averaging the performance metrics defined in Section II.A (horizontal speed, rate of descent, pitch angle, and pitch rate at touchdown) over the over the last 10 runs completed by each participant in each phase. These metrics are shown in Fig. 10 as box-whiskers plots to compare the performance of the two groups in the Training, Transfer and Back-Transfer phases.

Fig. 10a and 10c show the distribution of the average horizontal speed $\overline{V_{xt}}$ and of the average pitch angle $\overline{\theta_{td}}$ at touchdown, respectively. Although the EHE group has a larger within group variability, the performances of the two groups for these two metrics are comparable in each experiment phase and are flat throughout the experiment. This is
Table 7  Dependent-samples t-test between the experiment phases. Bonferroni correction was not applied.

<table>
<thead>
<tr>
<th>Metric Group</th>
<th>Group</th>
<th>Metric</th>
<th>From</th>
<th>To</th>
<th>t-test</th>
<th>From</th>
<th>To</th>
<th>t-test</th>
<th>df</th>
<th>Sig. (2-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HEH</td>
<td>Training</td>
<td>HEH</td>
<td>Transfer</td>
<td>−3.813</td>
<td>6</td>
<td>0.009*</td>
<td>Transfer</td>
<td>Back-Transfer</td>
<td>0.236a</td>
</tr>
<tr>
<td></td>
<td>EHE</td>
<td></td>
<td></td>
<td></td>
<td>0.786a</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>HEH</td>
<td>Training</td>
<td>HEH</td>
<td>Transfer</td>
<td>2.429</td>
<td>6</td>
<td>0.050*</td>
<td>Transfer</td>
<td>Back-Transfer</td>
<td>0.309</td>
</tr>
<tr>
<td></td>
<td>EHE</td>
<td></td>
<td></td>
<td></td>
<td>0.000***</td>
<td>5.562</td>
<td>6</td>
<td>0.012*</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* Statistically significant (p ≤ 0.05) difference between compared samples.
** Statistically highly significant (p ≤ 0.001) difference between compared samples.
* At least one sample not normally distributed. Related-samples Wilcoxon signed-rank test was applied instead of paired-samples t-test.

confirmed by the repeated measures ANOVA tests performed on the average horizontal speed and on the average pitch angle at touchdown, that do not show any statistically significant effects (Tab. 6).

Fig. 10  Distribution of the average performance at touchdown for each group in each phase.
A completely different trend can be observed for the average rate of descent at touchdown $V_{zt_d}$, as shown in Fig. 10b. The HEH group exhibits an improvement from the hard (training phase) to the easy dynamics (transfer phase), whereas performance is unaffected going from the easy (transfer phase) to the hard dynamics (back-transfer phase). A similar situation is found for the EHE group, whose performance degrades from the easy (training phase) to the hard dynamics (transfer phase) and improves from the hard (transfer phase) to the easy dynamics (back-transfer phase). These behaviors denote a correlation between the average rate of descent at touchdown and the number of successful landings, which is confirmed by the repeated measures ANOVA test performed on the average rate of descent at touchdown, shown in Tab. 6 (statistically significant within-subjects effect: $F(2,24) = 4.789, p = 0.018$).

The presence of an overall significant effect in the complete data set of the average rate of descent was further investigated by performing t-tests on individual sets of samples. Tab. 7 summarizes the results of these tests, highlighting the presence of a significant difference for the average rate of descent $V_{zt_d}$ between the training and the transfer phase for the HEH group (from the hard to the easy helicopter dynamics: $Z = -2.028, p = 0.028$) and from the transfer to the back-transfer phase for the EHE group (from the hard to the easy helicopter dynamics: $Z = -2.197, p = 0.043$).

<table>
<thead>
<tr>
<th>Table 8 Independent-samples t-test between the two groups.</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Metric</strong></td>
</tr>
<tr>
<td>$q_{td}$</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>$\Delta t_{rec}$</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>$h_{cush}$</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td></td>
</tr>
</tbody>
</table>

* Statistically significant ($p \leq 0.05$) difference between compared samples.

At least one sample not normally distributed. Independent-samples Mann-Whitney U test was applied instead of independent-samples t-test.

The surprising result concerns the average pitch rate $\bar{q}_{td}$, which is considerably different between the two groups during the training and the transfer phases, as shown in Fig. 10d. This is confirmed by a statistically significant between-subjects effect in the repeated measures ANOVA test of Tab. 6 ($F(1,12) = 6.237, p = 0.028$). The presence of an overall significant effect in the complete data set of the average pitch rate was further investigated by performing t-tests on individual sets of samples. Tab. 8 illustrates the results of these tests and highlights a statistically significant difference between the two groups during the training ($t(12) = -2.465, p = 0.030$) and the transfer ($t(12) = -2.245, p = 0.045$) phases.

It appears that the two groups adopt a completely different control strategy during the first two phases of the experiment: whereas the HEH group tends to touch down with a negative pitch rate (nose-down), the EHE group shows
a positive one (nose-up). The former behavior is usually adopted in reality in order to level the skids with the ground to avoid tail strike and have a better visibility before cushioning the touchdown [8]. The EHE group aligned with the HEH group during the back-transfer phase. In order to gain more insight into this unexpected result, a detailed analysis of the control techniques adopted by the pilots of the two groups is conducted next.

B. Control Strategy Metrics

As for the performance metrics, the control strategy metrics (reaction time, recovery time, flare initiation altitude, rotation altitude, cushion altitude, and rotor RPM at touchdown) were also averaged over the last 10 runs completed by each participant in each phase. These averaged metrics are shown in Fig. 11 as boxwhiskers plots to compare the control strategy of the two groups in the Training, Transfer and Back-Transfer phases. From Fig. 11, it appears that the spread of results for the EHE is generally larger than the HEH group, regardless of training, transfer or back transfer phase. The source of the larger spread for the EHE group compared to the HEH group is probably related to the fact that the EHE group indeed seem to use more variable strategies for attaining desired performance (e.g., anticipate the flare, cushion the touchdown before leveling the skids).

Both the average reaction time $\Delta t_{\text{reac}}$ and the average recovery time $\Delta t_{\text{rec}}$ are considerably different between the two groups during the training and the back-transfer phases, as shown in Fig. 11a and 11b, respectively. However, only for the recovery time this is confirmed by a statistically significant between-subjects effect in the repeated measures ANOVA test of Tab. 6 ($\Delta t_{\text{rec}}$: $F(1, 12) = 9.753, p = 0.009$).

Almost every participant of both groups is able to keep the average reaction time below 1 s, which is usually the value considered as pilot time delay following power failure during the certification of a civil helicopter [59]. Although the failure was random and unannounced, participants were expecting it to happen, keeping a high level of alertness. This might be the reason for such a good result in terms of reaction time. However, the participants of the HEH group react faster than those of the EHE group during the training and the back-transfer phases. This is explained by the fact that, during these phases, the HEH group was dealing with the hard dynamics, which inherently require a faster reaction than the easy dynamics, that give more margin for error or grants more spare capacity to the pilot. As expected, a faster reaction also leads to a faster recovery, as shown in Fig. 11b.

The presence of an overall significant effect in the complete data set of the average recovery time was further investigated by performing t-tests on individual sets of samples. Tab. 8 illustrates the results of these tests and highlights a statistically significant difference between the two groups during the training ($t(12) = -2.771, p = 0.017$ for the average recovery time) and the back-transfer ($t(12) = -3.108, p = 0.009$ for the average recovery time) phases.

Fig. 11c shows the distribution of the average flare initiation altitude $h_{fi}$, which is comparable for the two groups in each experiment phase and is approximately constant at around 120 ft throughout the experiment. This is confirmed by the repeated measures ANOVA test performed on the average flare initiation altitude, which does not show any
Fig. 11  Distribution of the average control strategy adopted by each group in each phase.
statistically significant effects (Tab. 6).

Some participants of the EHE group start to level the helicopter with the ground much earlier than the participants of the HEH group during the training and the transfer phases (Fig. 11d). This causes the rotor RPM to drop below 100% before pulling up the collective lever and leads to lower values of the rotor speed at touchdown (Fig. 11f). Lower RPM may result in a less effective cushion [8]. However, these differences between the groups in the average rotation altitude $\bar{h}_{rot}$ and in the average rotor speed at touchdown $\Omega_{td}$ are not significant and do not change significantly throughout the experiment, as summarized in Tab. 6.

For the average cushion altitude $\bar{h}_{cush}$ (Fig. 11e), a statistically significant between-subjects effect is found ($F(1, 12) = 5.158, p = 0.042$), as shown in Tab. 6. Indeed, in every experiment phase, the participants of the EHE group tend to start the cushion approximately 7 ft earlier than those of the HEH group. As can be verified from Tab. 8, only during the training phase a statistically significant difference between the two groups is found ($t(12) = -2.311, p = 0.039$).

Overall, the differences in the control strategy of the two groups, as shown in Fig. 11, can be used to motivate the difference in the average pitch rate at touchdown during the training and the transfer phases (Fig. 10d). Indeed, many participants of the EHE group touch down with a positive pitch rate in most of the runs of those two phases. This is mainly due to the fact that they pull up the collective to cushion the touchdown too early, resulting in a bounce, i.e., the helicopter gains altitude before touchdown. As a consequence, a considerable amount of rotor energy is dissipated and the loss of collective effectiveness is counteracted by starting a second flare.

**IV. Discussion**

The quasi-Transfer-of-Training experiment presented in this paper was designed to investigate how helicopter dynamics affect pilots’ acquisition of skills during autorotation training in a flight simulator. Two sets of helicopter dynamics, characterized by a different autorotative index (hard, lower index, and easy, higher index) [48], and two groups of participants, both chosen among experienced helicopter pilots, were considered. In order to assess whether familiarity with one set of helicopter dynamics affects the learning of new helicopter dynamics, each group started the training with either the hard or the easy dynamics, was then transferred to the other, and, finally, transferred back to the initial dynamics.

Although designed for a different training task, the outcome of this experiment confirms previous experimental evidence which showed positive transfer of skills from agile (hard case, where high compensation is required to the pilot) to inert (easy case, where low intervention is required to the pilot) dynamics, but not the opposite [36]. Indeed, in our experiment, both groups of participants exhibit a decrease in the rate of descent at touchdown from the hard to the easy dynamics, but not after a transition from the easy to the hard dynamics. This result corroborates our hypothesis, because a lower rate of descent is an indicator of more controlled and smoother touchdowns. The previous statement is
also supported by an increase in the number of landings within desired performance during the transfer from the hard to the easy helicopter dynamics of around 27% for the HEH group (significant effect) and 8% for the EHE group, which however was not statistically significant.

This is in agreement with current flight education, which usually starts with unaugmented helicopters at the beginning. Once proficiency is reached, later training stages involve augmented helicopters [60].

Pilots’ comments yielded a number of interesting recommendations to improve the setup used in this experiment and, more in general, current simulator training, and help to interpret some of the obtained results. For example, participants of both groups were able to achieve adequate performance at touchdown in most of the landings with both sets of dynamics. However, they struggled to attain desired performance for both the horizontal speed and the rate of descent at touchdown. This is most likely due to poor or lack of visual cues. Indeed, the SIMONA Research Simulator was conceived for fixed-wing simulation and is not equipped with chin bubbles. As a consequence, pilots completely lose sight of the ground during the flare. Therefore, pilots either opt for a less effective flare, ending-up with a higher horizontal speed, or risk either to strike the tail on the ground or not to coordinate the cushion timely, resulting in a higher rate of descent at touchdown.

Besides chin bubbles, another issue related to lack of visual cues is depth perception, which is the human visual ability to perceive the world in three dimensions and sense the distance of an object. Depth perception arises from a variety of depth cues. Not all the depth cues are equally important to every flying task. The depth cues of linear perspective, texture and parallax gradients, apparent/familiar size and streaming are highly useful during landing tasks. In particular, texture gradients and apparent/familiar size are the two elements that can be improved in the current out-of-the-window scenery. Texture gradients cue is the ability to perceive fine details on nearby objects. Such details are not visible on faraway objects. For instance, many pilots commented that, in reality, they start the flare when they see “the grass becoming grass”, that is when they realize that the grass is no longer a green expanse and they are able to discern the blades of grass.

Apparent/familiar size cue is the ability to determine the absolute depth of an object by combining the fact that the further the object is from the observer, the smaller it appears with previous knowledge of the object’s size. Therefore, the use of a finer texture and adding more familiar objects to the scenery may improve depth perception and increase pilots acceptance of the simulation environment.

The hard helicopter dynamics foster the development of more robust and flexible flying skills, because pilots are required to react faster to perceptual changes. Indeed, participants of the HEH group adopted, from the start of the experiment, a control strategy similar to the one adopted in real helicopters, as opposed to the participants of the EHE group, who tend to underestimate the altitude during the first two phases of the experiment, thus preempting the cushion. This sometimes results in a bounce (the helicopter gains altitude before touchdown), causing the rotor speed to drop down and the consequent loss of collective effectiveness that is counteracted by starting a second flare. This is the
reason why the participants of the EHE group touch down with a positive pitch rate during the training and the transfer phases. However, they align their control strategy with that of the participants of the HEH group during the back-transfer phase (from the hard to the easy dynamics).

Faster reactions to perceptual changes can translate into a safety enhancement, because pilots trained in high resource demanding conditions are more likely to be able to handle emergencies like engine failures in the real world, where the actual situation may easily divert from the training scenario. Results are promising and represent a solid foundation to extend this study. A new experiment with a more complex flight dynamics model, which incorporates also the helicopter lateral-directional dynamics, will be conducted to obtain more evidence for the findings presented in this paper, which are based on a flight dynamics model with longitudinal dynamics only.

Furthermore, given the success of the Control Event Detection (CED) methodology in identifying the different phases of the autorotation maneuver, future work will be conducted to assess its application in other types of experiments and control tasks, such as shipboard helicopter operations.

V. Conclusions

A quasi-Transfer-of-Training experiment with fourteen experienced helicopter pilots was performed in TU Delft’s SIMONA Research Simulator (SRS) to compare the effects of helicopter dynamics characterized by a high autototative flare index (hard dynamics) and low index (easy) on autorotation training in a flight simulator. Participants were divided in two groups and trained to perform a straight-in autorotation maneuver controlling a four degrees of freedom non-linear helicopter model with 3-DOF symmetrical dynamics plus rotorspeed. Each group tested the two sets of dynamics in a different training order: hard-easy-hard (HEH group) and easy-hard-easy (EHE group). Results show a positive transfer of skills from the hard helicopter dynamics to the easy dynamics for both groups, with the average rate of descent at touchdown that decreases of 94 ft/min for the HEH group and of 129 ft/min for the EHE group. This corroborates earlier findings that the acquisition of robust flying skills is fostered by initiating training in the most challenging setting.

In addition, participants of the EHE group adopted a sub-optimal control technique during the final part of the maneuver. This is suggested by the different sign of the pitch rate at touchdown for the two groups during the first two phases of the experiment: the HEH group tends to touch down with a negative pitch rate (nose-down), whereas the EHE group shows a positive one (nose-up). The former behavior is usually adopted in reality in order to level the skids with the ground to avoid tail strike and have a better visibility before cushioning the touchdown. Dealing with the difficult dynamics helped the participants of the EHE group to align their control strategy with that of the participants of the HEH group.

Although more experiments are needed to confirm our findings, results suggest that simulator training for autorotation can best start with training in the most resource demanding condition. Difficult dynamics require rapid responses to perceptual changes, forcing pilots to develop more robust and adaptable flying skills. This can enhance helicopter safety.
as pilots will be better prepared to face unexpected events that may occur during actual flight.

**A. Selection of Helicopter Configurations**

Scaramuzzino et al. [19] compared the stability characteristics in autorotation of helicopters with a different autorotative flare index (AI) [48]. Although a high autorotation index is desirable, it was shown that stability in autorotation does not always improve when the index increases. Therefore, an additional step was required to isolate only two configurations among those studied by Scaramuzzino et al. [19], consisting in a pre-experiment with a test pilot.

Thirty-two configurations were studied (Fig. 12a). They were the result of the variation from the baseline values (Bo-105 helicopter) of four design parameters (main rotor blade chord $c$, main rotor radius $R$, main rotor speed $\Omega$, and helicopter weight $W$) which affect the value of the AI. This was demonstrated by deriving an approximated expression of the AI [19]. Different types of helicopters have been considered in Fig. 2 and all of them have an AI between 5 and 40 ft$^3$/lb. Therefore, each of the four design parameters affecting the AI was varied individually to obtain eight different values of the AI evenly spaced from 5 to 40 ft$^3$/lb.

Fig. 2 shows that values of the index above 30 ft$^3$/lb can only be achieved by single-engine helicopters, whereas values below 15 ft$^3$/lb are typical of large helicopters. Since the baseline helicopter (Bo-105) is a medium weight twin-engine helicopter, those configurations with an AI index below 15 ft$^3$/lb and above 30 ft$^3$/lb were excluded. The configurations with an AI of 20 ft$^3$/lb and 25 ft$^3$/lb were also excluded because too close to the baseline, meaning that probably the pilot would not be able to notice any difference between these configurations and the baseline. Only the configuration with an AI of 20 ft$^3$/lb due to a main rotor blade chord ($c$) variation with respect to the baseline value was retained. Indeed, the main rotor blade chord is the only linear term in the approximated expression of the AI [19]. Therefore, this parameter is subject to the highest variations with respect to the baseline value and might lead to non-physically feasible configurations for high and low AI. This selection process resulted in nine configurations (Fig. 12b and Tab. 9) over the thirty-two initially available (Fig. 12a).

**Table 9 Configuration test matrix.**

<table>
<thead>
<tr>
<th>Design parameter</th>
<th>Autorotative flare index AI (ft$^3$/lb)</th>
<th>Blade chord $c$ (ft)</th>
<th>Main rotor radius $R$ (ft)</th>
<th>Main rotor speed $\Omega$ (rad/s)</th>
<th>Helicopter weight $W$ (lb)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blade chord</td>
<td>15</td>
<td>0.5692</td>
<td>0.1735</td>
<td>16.1</td>
<td>44.4</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>0.7589</td>
<td>0.2313</td>
<td>14.7</td>
<td>44.4</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>1.1385</td>
<td>0.3470</td>
<td>16.9</td>
<td>44.4</td>
</tr>
<tr>
<td>Main rotor radius</td>
<td>15</td>
<td>0.8858</td>
<td>0.2700</td>
<td>14.7</td>
<td>44.4</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>0.8858</td>
<td>0.2700</td>
<td>16.9</td>
<td>44.4</td>
</tr>
<tr>
<td>Main rotor speed</td>
<td>15</td>
<td>0.8858</td>
<td>0.2700</td>
<td>16.1</td>
<td>35.6</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>0.8858</td>
<td>0.2700</td>
<td>16.9</td>
<td>50.3</td>
</tr>
<tr>
<td>Helicopter weight</td>
<td>15</td>
<td>0.8858</td>
<td>0.2700</td>
<td>16.1</td>
<td>44.4</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>0.8858</td>
<td>0.2700</td>
<td>16.9</td>
<td>4279</td>
</tr>
</tbody>
</table>
A. Pre-experiment

To determine the easiest and the hardest configurations, a pre-experiment with a helicopter test pilot was conducted. The test pilot performed five autorotative landings with each of the ten pre-selected configurations, the baseline plus the nine listed in Tab. 9 and shown in Fig. 12b. During each run, pilot’s inputs and the helicopter model response to these inputs were collected. After the completion of each configuration (five runs), the pilot provided Cooper-Harper ratings (HQR) [61] according to performance defined in Tab. 1.

The results of the pre-experiment are shown in Fig. 13 for each of the four metrics considered in the performance standards of Tab. 1. Every box plot represents the variability in performance over the five autorotative landings performed with a specific configuration. Painted stripes are adopted to provide the reader with a visual aid to quickly assess pilot’s performance in every configuration. From Fig. 13c and 13d, it can be noticed that the pilot can easily attain desired performance for the pitch angle and the pitch rate at touchdown in every run with every configuration. The rate of descent at touchdown always falls within adequate performance (Fig. 13b), whereas the horizontal speed at touchdown falls outside performance standards in few runs with some configurations (Fig. 13a). Since the performance standards in terms of pitch angle and pitch rate were easily attained, the assessment of the different configurations was based only on the horizontal speed and on the rate of descent (Fig. 13a and 13b).

The configuration AI15c shows the highest variability and is characterized by the worst performance for both metrics. The opposite holds for the configuration AI30W. These results are in strong agreement with the subjective ratings provided by the pilot listed in Tab. 10. AI30W was chosen as the easiest configuration. AI15c, however, was not
selected as the hardest configuration, because it involves a 36% variation of the main rotor blade chord with respect to the baseline value, hence it was considered not physically feasible. Looking at both the objective metrics (Fig. 13a and 13b) and the subjective ratings (Tab. 10), the baseline configuration takes the second place as worst configuration and so can be safely regarded as the hardest configuration. Therefore, the A130W configuration, taken as the easiest configuration, represents a lighter version of the baseline helicopter that is instead taken as the hardest configuration. These two configurations were compared in the quasi-Transfer-of-Training experiment described in Section II to check whether one of them fosters the learning process of the autorotation maneuver.

![Desired performance](image1)

![Adequate performance](image2)

![Outside standards](image3)

**Fig. 13** Distribution of the performance at touchdown for each of the ten configurations.
Table 10  Cooper-Harper ratings (HQR) for the ten configurations flown by the test pilot.

<table>
<thead>
<tr>
<th>Configuration</th>
<th>Autorotation Index (ft^3/lb)</th>
<th>Parameter</th>
<th>Handling Qualities Rating (-)</th>
<th>Selected</th>
</tr>
</thead>
<tbody>
<tr>
<td>AI_{baseline}</td>
<td>23.3413</td>
<td>Baseline</td>
<td>5</td>
<td>Yes (Hard)</td>
</tr>
<tr>
<td>AI15c</td>
<td>15</td>
<td>Chord (c)</td>
<td>6</td>
<td>No</td>
</tr>
<tr>
<td>AI20c</td>
<td>20</td>
<td>Chord (c)</td>
<td>4</td>
<td>No</td>
</tr>
<tr>
<td>AI30c</td>
<td>30</td>
<td>Chord (c)</td>
<td>4</td>
<td>No</td>
</tr>
<tr>
<td>AI15Ω</td>
<td>15</td>
<td>RPM (Ω)</td>
<td>5</td>
<td>No</td>
</tr>
<tr>
<td>AI15Ω</td>
<td>30</td>
<td>RPM (Ω)</td>
<td>4</td>
<td>No</td>
</tr>
<tr>
<td>AI15R</td>
<td>15</td>
<td>Radius (R)</td>
<td>3*</td>
<td>No</td>
</tr>
<tr>
<td>AI30R</td>
<td>30</td>
<td>Radius (R)</td>
<td>4</td>
<td>No</td>
</tr>
<tr>
<td>AI15W</td>
<td>15</td>
<td>Weight (W)</td>
<td>5</td>
<td>No</td>
</tr>
<tr>
<td>AI30W</td>
<td>30</td>
<td>Weight (W)</td>
<td>3</td>
<td>Yes (Easy)</td>
</tr>
</tbody>
</table>

* Pilot comment: “This rating might be affected by learning effects”.

Acknowledgements

We thank the 14 participants of our experiment for their efforts. 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. This project has received fundings from the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska-Curie grant agreement No 721920.

References


