Human-automation interaction for helicopter flight
Comparing two decision-support systems for navigation tasks

Friesen, Daniel; Borst, Clark; Pavel, Marilena D.; Masarati, Pierangelo; Mulder, Max

DOI
10.1016/j.ast.2022.107719

Publication date
2022

Document Version
Final published version

Published in
Aerospace Science and Technology

Citation (APA)
Friesen, D., Borst, C., Pavel, M. D., Masarati, P., & Mulder, M. (2022). Human-automation interaction for helicopter flight: Comparing two decision-support systems for navigation tasks. *Aerospace Science and Technology*, *129*, Article 107719. https://doi.org/10.1016/j.ast.2022.107719

Aerospace Science and Technology 129 (2022) 107719

Contents lists available at ScienceDirect

Aerospace Science and Technology

www.elsevier.com/locate/aescte

Human-automation interaction for helicopter flight: Comparing two decision-support systems for navigation tasks

Daniel Friesen [a,b,*], Clark Borst [a], Marilena D. Pavel [a], Pierangelo Masarati [b], Max Mulder [a]

[a] *Delft University of Technology, Kluyverweg 1, 2629HS, Delft, the Netherlands*
[b] *Politecnico di Milano, Via La Masa 34, 20156, Milan, Italy*

ARTICLE INFO

Article history:
Received 20 July 2021
Received in revised form 9 June 2022
Accepted 15 June 2022
Available online 20 June 2022

Keywords:
Helicopter
Automation
Display
Human-machine interface

ABSTRACT

This paper investigates the effects of different automation design philosophies for a helicopter navigation task. A baseline navigation display is compared with two more advanced systems: an advisory display, which provides a discrete trajectory suggestion; and a constraint-based display, which provides information about the set of possible trajectory solutions. The results of a human-in-the-loop experiment with eight pilot participants show a significant negative impact of the advisory display on pilot trajectory decision-making: out of the 16 encountered off-nominal situations across the experiment, only 6 were solved optimally. The baseline and constraint-based display both lead to better decisions, with 14 out of 16 being optimal. However, pilots still preferred the advisory display, in particular in off-nominal situations. These results highlight that even when a support system is preferred by pilots, it can have strong inadvertent negative effects on their decision-making.

1. Introduction

Automation often works best in clearly defined tasks. However, during unanticipated or emergency situations [1], or when many automation systems are available to the pilots in parallel [2], automation can actually increase the requirements on the system operator or pilot. In unfortunate circumstances, automation can even contribute to accidents, as has happened in the fixed-wing domain [3]. While automation is certainly not the sole cause of accidents, unexpected automation reactions to unknown failures can exacerbate the dependency on pilot judgement in these situations. It is therefore required to scrutinise possible automation designs for both their possible positive and negative effects.

This paper investigates automation support for the cognitive task of navigational decision-making. The goal of this paper is to analyse what kind of automation system best supports this pilot task. Based on this analysis, recommendations for future helicopter automation design are derived. To that end, this paper compares two different automation design philosophies, advisory automation support and constraint-based automation support.

Advisory automation focuses on a clearly defined task and provides one particular solution to it. This solution (e.g., a specific manoeuvre, flight profile, or control strategy) is either communicated to the pilot or automatically implemented. Constraint-based automation takes inspiration from Ecological Interface Design (EID). Ecological interfaces provide information about the controlled system and its environment such that the constraints on possible operator actions become easily apparent [4]. Crucially, it leaves the decision-making task to the pilot. Visualised constraints can be physical (e.g., avoiding flight into terrain or bad weather) or procedural (e.g., staying above a predetermined safe altitude) [5]. EID principles have been only sparsely applied in the helicopter domain, for example for shipboard landing [6] and obstacle avoidance [7].

These two automation design philosophies manifest themselves in the three different helicopter head-down navigation displays that this paper investigates. A baseline display serves as an experimental baseline. It only shows the most necessary information about the position of the helicopter, the target, and any navigational obstacles. The first experimental display is based on advisory automation and provides one particular navigational solution to the pilots. This solution circumnavigates obstacles and provides a trajectory to the target. The second experimental display is based on Ecological Interface Design principles and provides information about the helicopter's navigational capabilities and limitations, without prescribing one specific solution. The inspiration for this constraint-based display lies in a display that was originally developed in the context of air traffic control [8], based on in-flight trajectory modification concepts developed by Mulder et al. [9].

---

[*] Corresponding author at: Delft University of Technology, Section Control & Simulation, Delft, the Netherlands.
E-mail address: d.friesen@tudelft.nl (D. Friesen).

https://doi.org/10.1016/j.ast.2022.107719
1270-9638/

Automated trajectory generation algorithms can rely on many different data sources: obstacle databases [10], the fuel cost of prospective trajectories [11], or the acoustic footprint of prospective trajectories [12]. Without going into detail, none of the aforementioned algorithms take all existing data into account. Rather, they focus on specific data subsets, relevant to the mission. That means that even if a trajectory determination algorithm considers all data that are deemed relevant for the mission, there is always the chance that other influences outside of the envisioned operational envelope require a manual change of trajectory. Detecting the departure from the operational envelope of the automation system is the responsibility of the pilots — the automation system is unable to react to data it is not programmed to consider.

This analysis, and the operational boundaries that all automation systems possess, highlight the crucial role of the pilots and their capability of adaptively reacting to situations. When the encountered situation lies outside of the scope of the automated system, or if an error prohibits the automation from working correctly, it is the pilots' responsibility to react to the situation and ensure the continued safety of the vehicle and the environment. This paper aims to provide insight into whether the use of the outlined automation design philosophies can support or hinder these adaptive pilot decision-making processes.

It would be unfeasible to try to design and evaluate systems that try to incorporate all the different kinds of data listed above or faults that may occur. However, it is also not necessary, as every automation system will have a specific operational envelope and can encounter situations outside of this boundary. The experiment of this paper reproduces and analyses this key characteristic: the experimental displays are designed within a particular operational envelope. They are subsequently subjected to situations inside and outside of this envelope. This enables the analysis of the pilots' reactions to both expected and unanticipated situations in a clearly defined context and how different automation design philosophies affect the pilots' decision-making.

This paper is structured as follows. Section 2 contains background information pertaining to helicopter navigation support systems. Section 3 describes the baseline navigation display and both experimental displays. Afterwards, an analysis of possible control strategies is performed in Section 4. The experimental setup is described in Section 5. The experiments' results are presented in Section 6 and discussed in Section 7, including recommendations for future research and automation design. Section 8 provides a conclusion to this paper.

2. Background

Top-down navigation displays are part of those electronic flight instrument systems that belong to second generation flight decks, which were introduced on a large scale with the Airbus A320 and the Boeing 747-400 [13]. On a navigation display, a multitude of information can be displayed, for example terrain and traffic data [13], heliport/heliport locations, restricted airspace and waypoints [14], or weather and obstacle data [15]. Coupled with a flight management system, a navigation display can provide information about waypoints and courses selected by the pilots [13]. Helicopter flight management systems, in particular, can offer mission-specific functions like automated flight pattern generation or the up- and down-link of flight plans with external sources [13].

Past work on helicopter navigation support systems includes the work of Haisch et al. [15], who describe the functionality of an envisioned adaptive route-planning algorithm. At the press of a button, a route from the current position to the mission target is calculated, taking into account data covering terrain, obstacles, topography, aerodromes, airspace, navigation, weather, and helicopter performance. When the system detects an additional obstacle, for example an additional bad weather area, the course is modified to evade the new obstruction. The calculated courses of this system seem to be made up of multiple straight legs between a small number of waypoints, i.e., no curved trajectories are proposed. The pilot can accept the proposed plan and "activate" it, or disregard it and insert a manual course with a joystick in the interseat console.

Takahashi et al. [16] performed an experiment that is similar to the one proposed in this paper. They investigated three different levels of automation support while performing a mission: fully coupled autonomy, additive control, and piloted decoupled attitude command. Trajectories to selected waypoints or landing sites were computed taking into account real-time obstacle data, and could either directly be implemented by the autopilot system or communicated to the pilot via head-down, panel-mounted displays. The focus of the experimental validation lied on the vehicle behaviour during mode transition and the manual control of the aircraft in the different modes.

Some current developments in pilot navigation support focus on the use of devices that are separated from the main avionics system of the helicopter, so-called Electronic Flight Bags (EFB). De Bernardi and Ferroni [17] describe Leonardo's EFB system Skyflight Mobile. The pilots are being made aware of possible terrain collisions, restricted airspace, warnings, and restrictions prior to the mission. Roos and de Reus [18] analyse the use of tablet-based EFBs in the helicopter cockpit, especially investigating effects on flight- and mission-safety induced by an additional, feature-rich tool in the cockpit. Based on available task load restrictions, the tablet could only be used in low-workload mission phases like cruise or holding. EFBs are not applicable to the continuous single-pilot operation envisioned in this paper.

The functionalities of the experimental displays in this experiment are not new. They are based on the extensive capabilities of currently operational systems, or those under research. However, by comparing different automation approaches, this paper will provide a deeper understanding of the interaction between automation and pilots, and what the impact of these interactions are.

3. Display design

3.1. Operational envelope

The operational envelope of all described displays is defined as the completion of a predetermined flight-plan, taking into account to-be-avoided weather areas, fuel constraints, and the track distance of the chosen trajectories. The goal of the navigation display is to support the following operator goals:

1. Perform a predetermined flight plan, which includes flying to the target waypoints in the order defined in the flight plan and hovering at each waypoint for ten seconds.
2. For each part of the flight plan, determine and execute a trajectory that is safe (i.e., does not enter weather areas) and efficient (i.e., in the constraints of this experiment, uses the path with the shortest track distance).
3. Provide predictions about the fuel use of the remaining flight plan legs.

The task of the pilots can be separated into two categories: 1) the manual flying task, which comprises hovering at each target and following the selected waypoints while avoiding bad weather areas; 2) the cognitive planning task, which comprises the selection of a suitable route to the next target and the evaluation of the remaining fuel with respect to the remaining legs.

A constraint is placed on the complexity of paths that will be supported by the experimental displays: only so-called "one-turn" trajectories between targets are supported. This means that

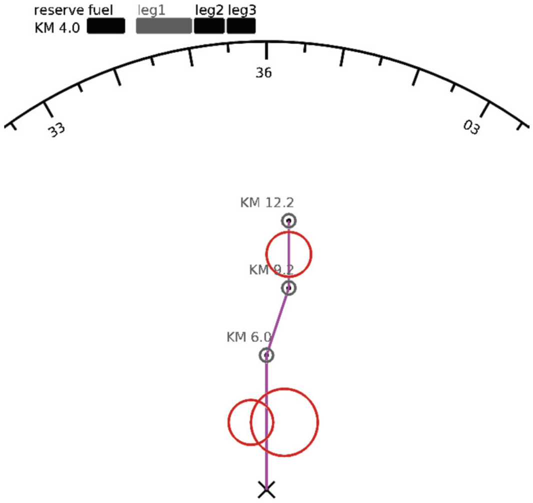

Fig. 1. Baseline navigation display representation of flight plan, including obstacles (red circles) and three targets (grey circles). The colours and line widths of the display have been adapted for better readability. (For interpretation of the colours in the figure(s), the reader is referred to the web version of this article.)

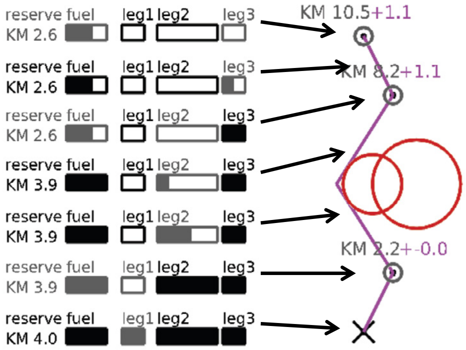

Fig. 2. Left: fuel gauge for reserve fuel and fuel assigned to legs one, two, and three, shown at different stages of an example mission. Right: one-turn trajectory to complete the example mission.

at most one intermediate waypoint is placed between the ownship position and the next target. Each trajectory between two targets also contains at most one intermediate waypoint. When following this kind of trajectory, the pilots need to perform at most one intermediate turn per flight leg. Fig. 2 (right) shows an example of a one-turn trajectory, containing zero turns between the ownship position and first target, one turn between the first and second target, and zero turns between the second and the third target.

This particular constraint on path complexity is chosen for a reason. A path with one intermediate waypoint is the first logical step between the most simple, direct path and more complex paths. Further steps to increase complexity would encompass increasing the number of intermediate waypoints, including curved/non-straight segments, and introducing time and altitude constraints. In this experiment, the border of the operational envelope is placed on the lower end of path complexity. This is done to enable the participating pilots to quickly understand the system boundaries and to swiftly learn to identify more complex path solutions outside of these boundaries.

Increasing the operational envelope to more complex trajectories would require the pilots to "think outside of an increasingly larger box". While this might increase the experiment realism, it would also require substantially more training and familiarisation with the proposed automation systems and the scenario. In addition, if it can be shown that certain automation systems have inadvertent negative effects in a straightforward navigation scenario where more complex trajectories are easily conceivable, it can be assumed that these negative effects are only exacerbated in more complex scenarios with less simple solutions and unclear system boundaries.

3.2. Baseline display symbology

The baseline navigation display utilised in this experiment shows a top-down representation of the outside world, see Fig. 1. The ownship aircraft is shown at its bottom edge. Target waypoints and bad weather areas are shown as small grey circles and larger red circles, respectively. Fig. 3 shows the rendering of the outside world as it is shown to the pilots in the simulator, depicting target waypoints one (red ground markings at the bottom of the figure) and two (red ground markings in the distance, only visible on the left-hand side of the figure) of the utilised example experiment course.[1] The current leg, which comprises reaching the next target from the ownship position, is called "active leg".

By pressing the "initialise" button on the cyclic stick, the pilots can trigger the calculation of the distance between themselves and the next waypoints (ignoring any obstacles which might be in the way). This distance is then shown next to the targets in the navigation display. The shortest, direct path from the ownship position to the remaining target waypoints is presented in magenta.

To identify a safe and efficient route, the pilots can only use the provided spatial information about the ownship position, the target position, and the position of any obstacles. The displayed direct route and distances can serve as a basis for the fuel estimation task. However, any deviations from the shortest path are not considered by the display. The pilots are required to estimate the additional travel distance themselves.

3.3. Fuel display

In every display condition, a fuel gauge shows pilots the remaining fuel which is planned for each leg and the remaining reserve fuel, see the lowermost fuel gauge on the left-hand side of Fig. 2. The remaining fuel is measured in track kilometres. It is computed by initially defining all fuel reserves in terms of track distance and continuously subtracting the sum of flown trajectory track distance from it. This allows the direct comparison of available fuel reserves to navigational distances, which is required for the experimental task. This fuel reserve calculation method is chosen to simplify the experimental task for the participants. At the same time, it is sufficient to introduce track efficiency considerations into the experiment. It does not consider the impact of flying at different velocities, which would change the consumed fuel per distance flown, or the fuel consumption during hover.

During the mission, the fuel "container" that is currently being emptied is highlighted. The leg-specific container contains enough fuel to complete the leg without any deviations from the shortest path: every deviation due to obstacles requires the use of reserve fuel. When the leg-specific fuel container is empty, reserve fuel is consumed. Subsequently, the reserve fuel gauge is highlighted.

Fig. 2 depicts the fuel gauge at different times during an example mission, when following the shown trajectory. The reader is advised to start at the bottom of this figure. At the beginning, at the lowermost position, all containers are full. After performing the first leg (i.e., after flying from the initial position to the first target), the corresponding container is empty. As there were no major course deviations necessary, the reserve fuel container is almost full. After performing the second leg (i.e., flying from the first

---

[1] Bad weather areas do not have a graphical representation in the simulation. Rather, the default visibility of 1,500 m is gradually reduced to 100 m when entering an area designated as bad weather.

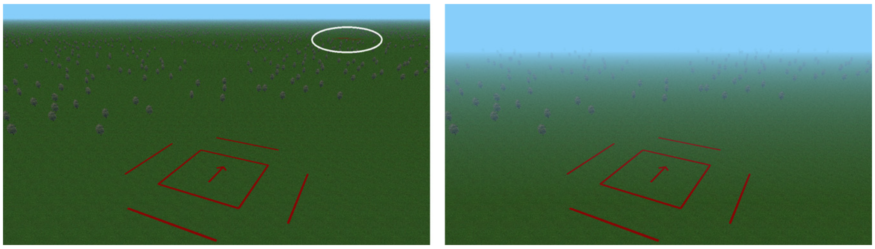

Fig. 3. Rendering of the experiment course, looking from target 1 red ground markings at the bottom of the figure to target 2. Left: increased visibility to highlight the second target waypoint. The oval is inserted to highlight the target and is not visible during the experiment. Right: actual visibility employed during the experiment (1,500 m).

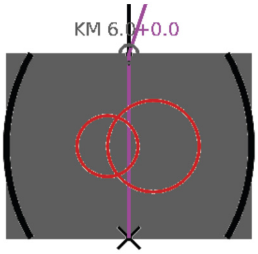

Fig. 4. Area covered by evaluation grid (201 times 101 points, grey); the black ellipse denotes the outermost possible locations of intermediate waypoint between ownship position (bottom centre) and target (top centre), given the remaining fuel.

to the second target), the second container is empty, and a non-negligible amount of reserve fuel has been used. This is caused by the fact that some deviation from the direct route was necessary to avoid entering bad weather areas. The remaining reserve fuel is displayed in kilometres on the left. At the end of the course, all three leg-specific containers are empty. In this example, the leg-specific fuel was sufficient to complete the last leg, with 2.6 km of reserve fuel remaining.

3.4. Trajectory determination and evaluation algorithm

Both experimental displays rely on an algorithm that determines all one-turn trajectories to the target and evaluates the determined trajectories with respect to safety (entry into bad weather) and efficiency (fuel consumption). When the calculation of these data is triggered, the area between the ownship position and the target position is first divided into a grid of 201 (lateral) times 101 (longitudinal) point locations. The lateral expansion of the grid is chosen such that it covers all possible one-turn trajectories between the current ownship position and the target. Fig. 4 shows the evaluation grid in grey.

Afterwards, for each point location, it is determined whether following a trajectory through this intermediate waypoint satisfies the safety requirement of not entering any weather areas. If the trajectory is safe, it is evaluated with respect to the length of the resulting trajectory. The length of the resulting trajectory is compared to the theoretically optimal, direct trajectory length. The shorter the trajectory is, the more efficient it is.

The result of this algorithm is a grid of location points, each with a binary safety value (safe/unsafe) and a numerical efficiency value (additional travel distance, compared to theoretical optimum). These data are used both by the advisory and constraint-based display, as described below.

3.5. Advisory display

The advisory display shows the same information as the baseline display. However, when the “initialise” button on the cyclic stick is pressed, the most efficient and safe intermediate waypoint of the previously computed grid is selected. Then, a trajectory is plotted from the ownship position, through the location point, to the target location. The resulting path and distances are then shown on the display, as is visible in Fig. 2.

The advisory display provides the pilots with a safe and optimal one-turn route to reach the target. The additional track distance required to follow the computed path is shown next to each target. This additional track distance directly relates to the remaining reserve fuel: when this trajectory is followed precisely, the reserve fuel will be reduced by the indicated amount when reaching the respective waypoint.

3.6. Constraint-based display

The constraint-based display provides pilots with graphical information about all possible collision-free, one-turn trajectories to reach the current target. It also shows the remaining manoeuvre capabilities for future legs, taking into account the remaining reserve fuel. As is shown in Fig. 5 on the left-hand side, multiple ellipses are shown around the direct flight path between the ownship position and the next target. Flying a one-turn trajectory with a turning point on the first ellipse results in an additional travel distance of one kilometre. Each following ellipse represents one additional kilometre of travel distance. Through the size of the ellipses, the pilots can estimate the additional distance that is required to complete the respective path. A highlighted area denotes the locations of all possible collision-free turning points in the active leg.

For the currently active leg, ellipses are shown at additional travel distances of 1 km, 2 km, 3 km, and 4 km. Within this area, the pilots can manually set a turning waypoint. By pressing the “select” button on the cyclic stick, the pilots can cycle through the ellipses of the current leg, as shown in Fig. 5 on the left. By turning the helicopter, the pilots can aim the nose of the helicopter at a certain point on the selected ellipse, see Fig. 5 in the middle. The pilots can select the intersection point of the ownship orientation and the selected ellipsis by holding the “select” button, as shown in Fig. 5 on the right. The distance to the currently active target is then re-calculated, considering the selected turning point. (Note that the distance between future target points is still the shortest direct distance, pilots can only manipulate the active leg.)

For future legs, only one ellipse is shown. The size of this ellipse depends on the remaining fuel reserve, reduced by the fuel requirements of the selected course in the current leg. As such, the constraint-based display supports the selection of a safe and optimal route to reach the current target. For subsequent targets, its support is weaker: it only shows the maximum extra track distance that the remaining fuel allows — intermediate ellipses are not shown.

4. Control strategy analysis

Fig. 6 depicts a Decision Ladder (DL), based on work by Rasmussen [19], for the path planning task. Possible control strategies afforded by the displays “move” through the DL, cover-

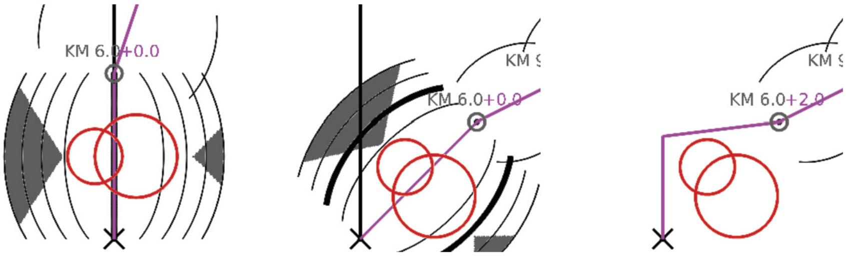

Fig. 5. Selecting a specific waypoint by interacting with the constraint-based display. Left: ellipses and area for safe turning point is shown. Middle: second ellipse is selected, helicopter is rotated to choose turning point. Right: turning point is selected.

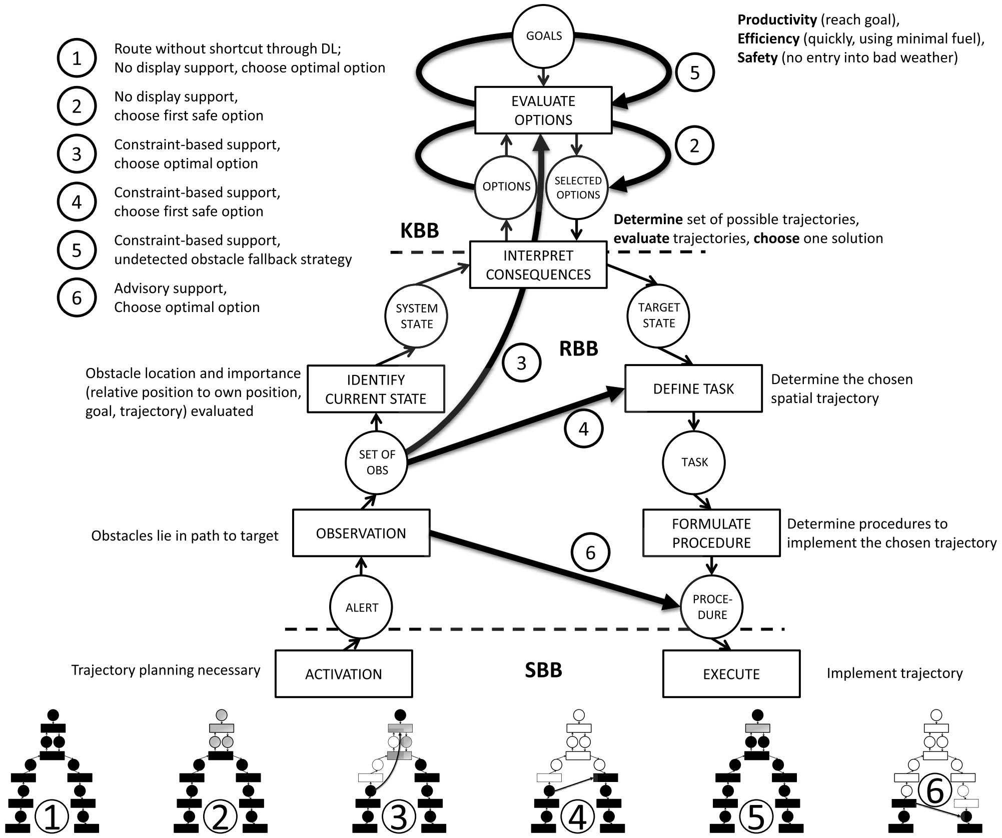

Fig. 6. Decision Ladder for the path planning task, showing the control strategy without shortcuts (1) and five control strategies with "automation–enabled shortcuts" (2) - (6).

ing Skill-Based Behaviour (SBB), Rule-Based Behaviour (RBB), and Knowledge-Based Behaviour (KBB).

4.1. Path planning strategies

Strategies (1) and (2) can be used regardless of the employed display. Strategy (1) comprises every step on the decision ladder, requiring knowledge-based reasoning and decision-making throughout the process. After initiating the path planning process "activation", the visible obstacles are identified, and their location and size are evaluated with respect to the ownship and target position. Then, possible solution paths are determined and evaluated. Based on the task-specific goals (safety, efficiency), one solution is chosen. This solution is then translated into an intermediate waypoint between the ownship position and the target, defining the selected path. Lastly, the chosen path needs to be implemented by the pilots by performing certain standard flying manoeuvres.

Strategy (2) is the first strategy that uses a rule-based "shortcut" in the DL. When evaluating possible solutions, the pilots might decide to choose the first viable route they encounter, neglecting part of the efficiency evaluation and only focusing on safety. The rule can be formulated as: "if a safe route is identified, then immediately implement this route and stop searching for alternative routes". In this case, the path determination, evaluation, and selection step is shortened, but at the possible expense of track efficiency.

Strategies (3) and (4) are enabled by the constraint-based display. With this display, after recognising the existence of an obstacle, the pilots can trigger the calculation and visualisation of all safe one-turn solutions. The pilots can immediately skip to a future step. In case of Strategy (3), they can use a "knowledge leap" to immediately skip to the path evaluation step: all safe possible one-turn paths are already calculated. The selection of the optimal path is supported through the ellipses, too, by visualising the additional track-distance of the possible turning points in one-kilometre increments. To choose the optimal path, the pilots need to determine the waypoint that is closest to the direct connection between the ownship position and the target, i.e., the safe waypoint that has the smallest additional track distance. After determining this waypoint, they can manually insert this waypoint into the navigation display.

The constraint-based display enables a second, larger shortcut, described in Strategy (4): instead of evaluating each proposed solution to choose the optimal route, the pilots can decide to choose the first available, safe solution. In this case, they choose and manually insert an arbitrary waypoint in the safe area. This rule-based shortcut will ensure a safe trajectory, but not an optimal one. It can be described by the following if-then clause: “if the constraint-based display provides any safe one-turn trajectories, arbitrarily select one solution and immediately implement it.”

Strategy (6), enabled by the advisory display, provides the largest rule-based shortcut. As soon as the pilots identify the need to perform the path-planning task, they trigger the automatic path planning system. This will automatically insert a safe and optimal one-turn waypoint into the navigation display. The pilots only need to implement the proposed route. This shortcut can be described as “if it is not possible to directly fly to the target, trigger the automatic path planning system and implement the suggested route.”

The advisory and constraint-based displays encourage certain control behaviours and shortcuts. Their impact on the decision-making process of the pilots depends on how prone pilots are to follow these shortcuts. How frequently do the pilots check the provided shortcuts for errors, and how frequently do they reflect on the requirements for the shortcuts to work? On the one hand, relieving the pilots of some cognitive work through shortcuts could lead to an increased mental capacity to evaluate and reflect on the current course of action. On the other hand, utilising shortcuts that skip the evaluation of the chosen trajectory by the pilots themselves could lead to a decrease of the level of scrutiny the suggested trajectories are subjected to.

The answers to these questions heavily depend on the mindset of the pilots. Are they expecting errors and unsafe system behaviour, or do they generally accept the provided shortcuts? In this experiment, while they were warned that additional obstacles might appear, it was not an emphasised element of their briefing. It can be reasonably assumed that they were mostly focused on the normal performance of the task, utilising the provided support, without questioning the provided automation support at every step of their thought process. This expectation is later translated into hypotheses.

4.2. Reacting to additional weather

To simulate situations that lie outside the operational boundary of the automation systems, obstacles can appear mid-run that remain undetected by the algorithm. This requires the pilots to detect this additional obstacle (and, when using any of the displays, the display malfunction) and perform all tasks themselves. To elaborate: the obstacles will still be shown on the navigation display if they affect the currently active leg, but both the automatic path calculation of the advisory display and the area of safe intermediate waypoints of the constraint-based display will be calculated without this particular obstacle. These events will enable the analysis of the robustness of the displays towards system malfunction. The right-hand side of Fig. 7 depicts an experiment course that contains an additional bad weather area that appears when entering leg 2. Fig. 8 shows how an undetected obstacle appears on the display when using the advisory display (left) and the constraint-based display (right).

The appearance of such an obstacle can cause a previously chosen trajectory to become unsafe. The realisation that a chosen trajectory is unsafe occurs during the “obstacle locations evaluated” step in the DL. Any strategy that provides shortcuts within or around this step is susceptible to undetected obstacles impairing the safety of the current leg. These are Strategies (3), (4), and (6).

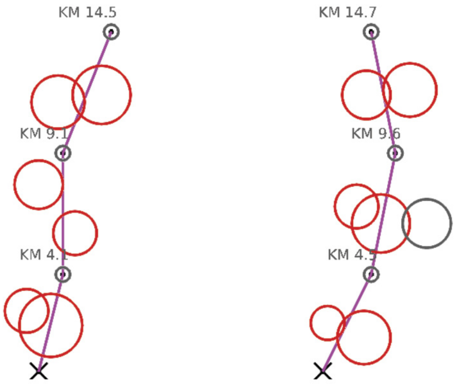

Fig. 7. Left: first experiment course design with a 2-turn solution at leg 2. Right: second experiment course design with an additional weather area appearing in leg 2 shown in grey.

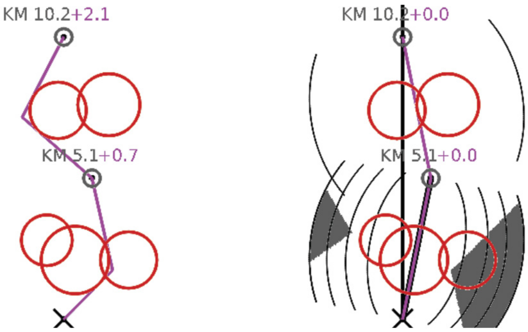

Fig. 8. Flawed display support by the advisory (left) and constraint-based (right) display when an additional weather area appears.

Strategies (1) and (2) can always be employed by the pilots. They are not susceptible to obstacles that appear mid-run. As long as the pilots recognise all present obstacles when they initiate the planning task, the path planning strategy and the safety and efficiency of the chosen path are not impaired.

Both Strategies (3) and (4) are vulnerable to undetected obstacles, as the calculated safe trajectories do not take this additional obstacle into account. The pilots are required to recognise the malfunction and manually adapt the suggested paths. As soon as the pilots recognise the additional obstacle, they have two options. First, they could decide to neglect the additional information of the constraint-based display completely and solely rely on the baseline data representation. In this case, they would change from Strategy (3) or (4) to the baseline Strategies (1) or (2). Second, they could decide to utilise those parts of the constraint-based display that are still valid, i.e., the additional track distance ellipses. This behaviour is represented as Strategy (5).

Strategy (5) is a fallback strategy for the constraint-based display, in case an undetected obstacle appears. In this case, the spatial representation of possible intermediate waypoints is no longer valid: some of the suggested trajectories will intersect the undetected obstacle. However, the pilots can still utilise the ellipses indicating additional track distances to manually evaluate a trajectory with respect to its additional fuel cost. This remaining display function can be represented as a shortcut within the “evaluate options” block in the DL.

The large shortcut of Strategy (6) is most susceptible to undetected obstacles, as pilots are not required to analyse and integrate the provided spatial information of the display in any way. To detect the error, they need to consciously analyse the proposed solution for any obstacle intersections. If the proposed solution is unsafe, there is no way of “fixing” this display error. Therefore, the

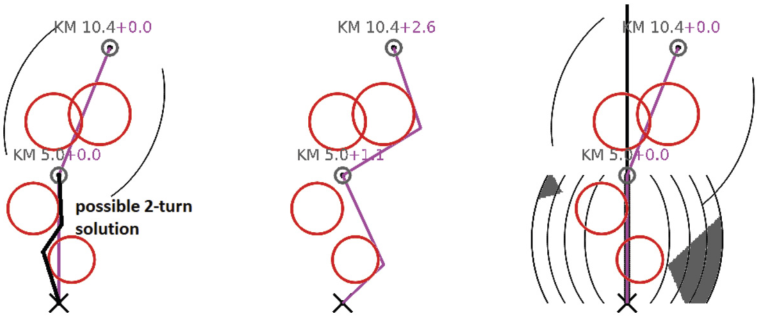

Fig. 9. A possible 2-turn solution (added in black), afforded by the gap between the obstacles, when viewed with the baseline (left), advisory (middle), or constraint-based (right) display. Both the advisory and constraint-based displays only suggest suboptimal routes around both weather areas.

Fig. 10. SIMONA Research Simulator at Delft University of Technology.

pilots need to disregard the display suggestions entirely and use either Strategy (1) or (2).

4.3. Insufficient fuel discovery strategies

The baseline display does not provide any support to estimate the additional distance that is necessary to avoid obstacles. The task of estimating this extra distance and connecting it to the remaining reserve fuel is left to the pilots.

The advisory display provides distance estimations for every target. If, in the planning phase, the additional distance is larger than the remaining reserve fuel, a completion of this trajectory is no longer possible. If an additional weather area appeared, the provided information is no longer accurate.

The constraint-based display provides support to discover insufficient fuel while planning the next leg. After selecting a trajectory in the current leg, the ellipses around future legs shrink to reflect the change in available reserve fuel (see Fig. 5). In this example, the ellipses in legs two and three do not intersect with weather areas and the course can be completed with the remaining reserve fuel. However, if the ellipses around future legs do not afford any trajectory solutions, i.e., if they would intersect with weather areas on both sides, the remaining reserve fuel will not be sufficient to complete the corresponding leg. In case of an additional appearing bad weather area, this information can still be used by the pilots.

5. Experimental setup

5.1. Scenario

The pilots are tasked to complete a predetermined flight plan which includes three target waypoints per experiment run, see Fig. 1 for an example. At each target, the pilots are asked to hover in place for ten seconds.

The path to the target waypoints (but not the target points themselves) can be obstructed by circular bad weather pockets (red), which must be evaded. The pilots are asked to approach each target waypoint as fast as possible and without entering bad weather areas, not exceeding a maximum speed of 100 kt.

During each experiment run, the helicopter possesses a certain amount of reserve fuel (measured in travel distance, four kilometres per run). If the pilots expect to run out of reserve fuel before reaching the next target, they need to detect this and abort the mission at the current position (by telling “mission control”, i.e., the experiment conductor).

It is of particular interest to investigate the effect of the employed displays on pilot decision-making during situations that do not neatly fall into the operational envelope of both displays (namely, the assumption that one-turn solutions are close to the optimal solution). To this end, two more complex obstacle arrangements are introduced.

In most cases, the one-turn solutions proposed by the displays are close to the optimal trajectory. However, depending on the location of the weather, there can be trajectories between waypoints that are more efficient than one-turn solutions, see Fig. 9. The suggested one-turn solutions lead completely around both obstacles, as shown by the advisory display (middle) and the constraint-based display (right). However, the most efficient route leads through the gap between the obstacles. This route is not detected by the support displays.

In addition to these more optimal 2-turn solutions, some additional bad weather pockets will appear for a small number of active legs and will not be recognised by the experimental displays. An appearing obstacle will always change the direction of the optimal trajectory: if, before the obstacle’s appearance, the shortest trajectory leads around the left side, the shortest path will afterwards lead around the right side and vice versa.

5.2. Apparatus

The experiment took place in March 2021 in the SIMONA Research Simulator (SRS) [20] at Delft University of Technology, shown in Fig. 10. The outside visuals with a field-of-view of 180° by 40° are collimated, appearing at an infinite distance to the pilots. The simulator windows resemble a fixed-wing cockpit, obstructing any downward view. For the given navigation task and the very large hover area this field-of-view limitation did not appear to play a detrimental role to the ability of the pilots to control the helicopter.

The participants used a helicopter cyclic stick, a collective stick, and pedals to control the model, which is an analytical model based on a Messerschmitt-Bölkow-Blohm Bo105 Helicopter [21]. The trigger of the cyclic stick served as the "initialise" button, a button close to the resting position of the right-hand thumb served as the "select" button. The motion system of the simulator was deactivated. The additional motion cues were expected to have a negligible influence on the cognitive task of decision-making in a navigation scenario. The added immersion was deemed insufficient to justify the added complexity and experiment duration which follows from the use of the motion system.

5.3. Participants

Eight helicopter pilots with varying experience (minimum Private Pilot License (PPL), approximately 100 flight hours) participated in this experiment. Five participants had a private helicopter pilot licence (PPLH), three participants had a commercial or more advanced helicopter licence. Average flight hours per participants amounted to 1,500 hours, with a standard deviation of 1,850 hours. Before the experiment, pilots could accustom themselves with the controls, the model, each experiment condition, and the experimental procedure.

5.4. Independent variables

The experiment utilised a within-participants design, each participant performed each condition. The independent variables of this experiment are *display* (baseline, advisory, constraint-based) and *situation* at the second leg of each course (more optimal 2-turn, undetected weather). These situations correspond to the two "more complex" obstacle arrangements described above. Each experimental course contains one of these obstacle arrangements, see Fig. 7.

Each course is flown with each display, resulting in six experimental runs per pilot. To avoid the recognition of the same course, the course elements are rotated between displays. This does not change the distances or obstacle location relative to the leg origins and targets. The experimental setup is therefore treated as a "within subject" design, even though there are technically six different courses. The order of experiment conditions is changed between pilots, to create a balanced experiment setup.

5.5. Dependent measures

Dependent measures comprise of decision-making, measured through the trajectory decision the pilots make; safety, measured via the number of "unsafe" fuel predictions (i.e., overestimating own capabilities); workload, measured via the subjective NASA-TLX, given to the pilots after each condition [22]; situation awareness, measured via the subjective scale Situation Awareness Rating Technique (SART) [23], likewise given to the pilot after each experiment condition; and pilot preference, measured through a questionnaire given to the pilots at the end of the experiment.

Decision-making and safety ratings are collected per leg and are analysed as such. Therefore, an experiment run always contains one data point for the first, nominal leg and one data point for the situation encountered at the second leg (2-turn possible or additional weather). The third leg is excluded, as some pilots were able to complete the third leg with the remaining fuel in some runs, but most pilots were not.

Workload and situation awareness ratings are collected per run. They therefore always contain at least one nominal leg and one situation at the second leg as the basis for the subjective rating. The comparative ratings for the NASA-TLX are only collected once per display, so three times in total. The weights are then applied to both runs with the same display.

5.6. Control variables

Control variables comprise the simulator setup, task, the utilised helicopter model, the baseline navigation display elements, and the instrument panel.

5.7. Data processing

Given the relatively small number of eight participants, only conservative, non-parametric test statistics are used. To compare numeric measures, non-parametric two-way Friedman tests or, when analysing data subsets with only one independent variable, one-way Kruskal-Wallis tests are employed. To compare binary measures, Cochran-Q tests, as implemented in MATLAB by Jos,[2] are utilised. All employed tests analyse the difference between multiple test attempts. A significant test result suggests that the observed differences are not based on random chance.

The data are treated on a "per course" basis. The course identifier is either C1, which is the experiment course with a possible 2-turn solution, or C2, which is the experiment course with an additional weather area appearing at leg 2.

Tests are performed at an initial significance value of $\alpha = 0.05$. The initial test takes all data of one course (either C1 or C2) into account. In this arrangement, the first independent test variable is display (baseline, advisory, constraint-based), and the second independent variable is leg number (leg 1, leg 2), which corresponds to nominal and off-nominal situations (again, either 2-turn possible or additional weather).

Post-hoc tests on subsets of the data are performed with a significance value of $\frac{\alpha}{n}$, where n is the number of subset tests. In most cases, n is equal to 5, when five subset tests are performed: three to analyse the effect of situation for each of the three separate displays and two to analyse the effect of display for each of the two situations. This Bonferroni-correction is carried out to achieve a significance value of $\alpha = 0.05$ for the combined post-hoc tests, accounting for the increased number of tests on the same data [24, p. 67]. Without this correction, the significance of the performed post-hoc tests would be overestimated.

5.8. Hypotheses

In nominal situations (legs without a possible 2-turn solution or additional weather), the advisory display will lead to the best trajectory decisions (i.e., go left or right around weather). The constraint-based display also enables good decision-making in these cases, but not as fast and direct as the advisory display.

The detection of two-turn solutions decreases when utilising the advisory display. The detection of these "unconventional" solutions takes place in the "determine, evaluate, select solution" block in the DL, Fig. 6, which is completely skipped with the advisory display in nominal situations, Strategy (6). With the baseline and constraint-based displays, the pilots are more involved with the spatial aspects of the prospective trajectories, which will lead them to detecting the two-turn solutions more often.

---

[2] Jos (10584) (2021). COCHRAN Q TEST (https://www.mathworks.com/matlabcentral/fileexchange/16753-cochran-q-test), MATLAB File Exchange. Retrieved March 29, 2021.

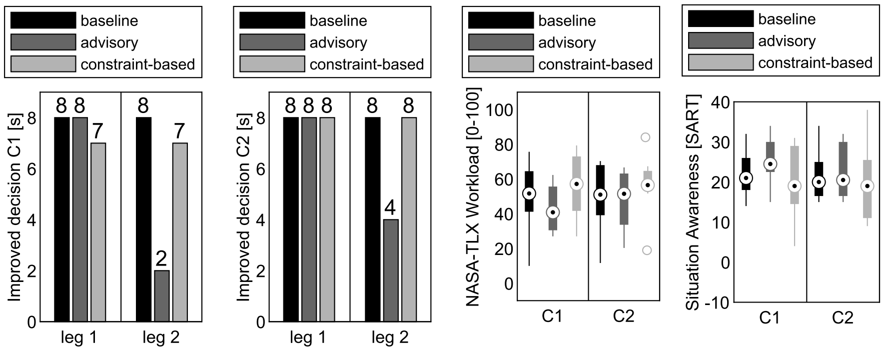

Fig. 11. Amount of optimal pilot decisions course C1 and course C2 (left), eight decisions per condition; subjective pilot ratings workload and situation awareness (right), a high rating indicates a large workload/good situation awareness, and vice versa.

During legs with additional bad weather areas appearing, the constraint-based display will lead to the best trajectory decisions, as parts of it can still be used to judge prospective trajectories according to DL Strategy (5). Both the advisory and baseline display will lead to worse decisions, as both displays can only rely on DL Strategies (1) or (2).

It is expected that pilots prefer the advisory display in nominal situations and the constraint-based display in off-nominal situations. This outcome would reflect results obtained in the fixed-wing domain [25].

6. Results

One pilot repeatedly hit the physical limits of the control inceptors, resulting in inconsistent helicopter model behaviour. This caused the participant to change the given fuel predictions. The results of this participant are therefore only included in the decision-making category, as this specific dependent measure is expected to be independent from the encountered model behaviour changes. The results of this pilot have been omitted in all other dependent measures. This behaviour was not caused by significantly lower or higher flight experience; the participant had comparable experience to other participants who did not cause this model behaviour.

6.1. Pilot decision-making

The number of optimal pilot decisions is shown in Fig. 11 (left). In this experiment, optimal pilot decisions are defined as follows: in case of possible 2-turn solutions, the optimal pilot decision is defined as "discovering" this hidden, more optimal solution and performing it, i.e., flying through the gap between the two weather areas. In case of additional weather, the optimal pilot decision is defined as choosing the shorter route around the weather area. All other trajectory decisions are defined as not optimal.

Considering course C1, there is a significant effect of display ($\chi^2(2) = 8$, $p < 0.05$) and of situation on pilot decision ($\chi^2(1) = 4.5$, $p < 0.05$). The number of optimal decisions when using the advisory display significantly drops when encountering the possible 2-turn solution at leg 2, corroborated by a significant effect of display in leg 2, $\chi^2(2) = 10.3333$, $p < 0.01$. Only 2/8 pilots chose the more optimal two-turn solution with the advisory display, compared to 8/8 with the baseline and 7/8 with the constraint-based display.

Analysing course C2 reveals a similar picture. Across all data, there is a significant effect of display ($\chi^2(2) = 8$, $p < 0.05$) and of situation ($\chi^2(1) = 4$, $p < 0.05$) on pilot decision. Analysing leg 2 separately reveals no significant effects of display ($\chi^2(2) = 8$, $p = 0.018$). However, there is a clear trend of worse decisions with the advisory display when encountering additional weather areas. Only 4/8 pilots chose the optimal route around the weather areas, the other pilots chose the less optimal direction suggested by the advisory display.

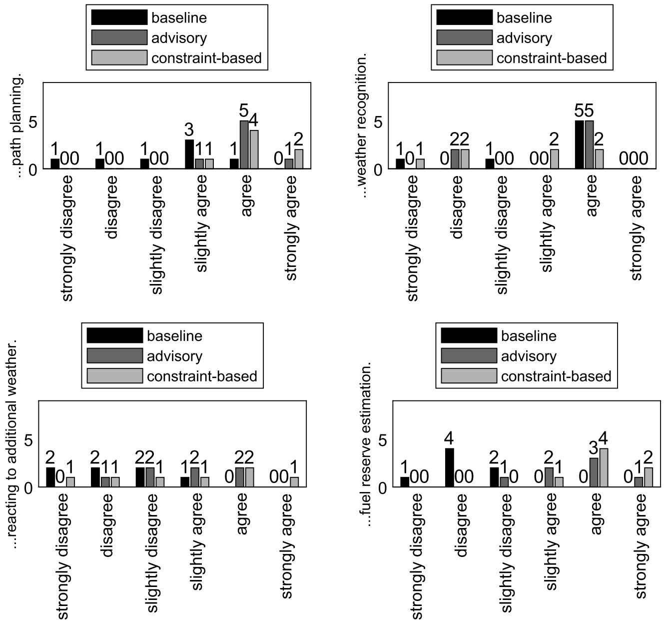

Fig. 12. Pilot opinion on aspects of display support: "The display supported...". The end of the respective question is shown to the left of each plot, possible answers are shown at its bottom. The number of answers in each category is shown on top of each coloured bar.

Fig. 12 shows the questionnaire results covering the perceived display support for the sub-tasks of path planning, weather recognition, and the reaction to additional weather areas. The only significant effect can be observed for the path planning task: the employed display significantly affected the answer to this question ($H(2) = 9.9884$, $p < 0.01$). Both the advisory and constraint-based display have higher ratings than the baseline display.

This result seems to contradict the previous results, which highlighted the negative effect of the advisory display on the quality of trajectory decisions. There are multiple possible explanations for this apparent discrepancy. First, the pilots might have rated the theoretical, abstract capability of the displays to support their path planning instead of the actual improvement they could observe during the experiment. It appears only logical that additional information, be it advisory or constraint-based, should have supported the pilots in planning prospective trajectories. This could explain why both displays have been rated higher than the baseline display.

A second possible explanation could be that the pilots did not rate the display support with regards to improved decision-making

Table 1
Pilot fuel over- and underestimations for Course C1, for the next leg and for the remaining course. Displays: baseline (B), advisory (A), constraint-based (C).

Leg Display	leg 1			leg 2			leg 3		
	B	A	C	B	A	C	B	A	C
overestimated leg capabilities	0	0	0	0	0	0	1	0	1
underestimated leg capabilities	0	0	0	0	0	0	1	0	0
leg estimation true	7	7	7	7	7	7	5	7	6
overestimated course capabilities	3	0	3	1	0	2	X	X	X
underestimated course capabilities	0	1	1	0	1	1	X	X	X
course estimation was true	4	6	3	6	6	4	X	X	X

Table 2
Pilot fuel over- and underestimations for Course C2, for the next leg and for the remaining course. Displays: baseline (B), advisory (A), constraint-based (C).

Leg Display	leg 1			leg 2			leg 3		
	B	A	C	B	A	C	B	A	C
overestimated leg capabilities	0	0	0	0	3	1	0	0	0
underestimated leg capabilities	0	0	0	1	0	0	0	0	0
leg estimation true	7	7	7	6	4	6	7	7	7
overestimated course capabilities	5	7	7	0	3	2	X	X	X
underestimated course capabilities	0	0	0	0	0	0	X	X	X
course estimation true	2	0	0	7	4	5	X	X	X

outcomes, but rather decision-making convenience. Both display variants provide shortcuts in the decision-making process that the baseline display does not offer. The pilots also might have been oblivious to the fact that their chosen trajectory was non-optimal, which can lead to the subjective perception that the provided display support was useful, even though it was not.

It seems like the advisory display has a detrimental effect on pilot decision making when encountering off-nominal situations. These worse decisions affected the safety values discussed in the later sections, as the chosen trajectories were less efficient and required more fuel than the optimal route.

6.2. Safety

Two different fuel predictions were made by the pilots. They needed to predict whether the remaining fuel is sufficient to complete the next leg, and they needed to predict whether the remaining fuel is sufficient to complete the whole course. Tables 1 and 2 show the number of over- and underestimations given by the pilots, separately per course, leg number, and display.

To determine the safety of the pilots' fuel predictions considering the next leg, three possible outcomes are considered. First, the pilots could overestimate their fuel capabilities. In this case, they predicted that they could finish the next leg, but they ran out of fuel before doing so. For course C1, no significant effects are observed, $p > 0.05$. Both the baseline and constraint-based display caused one overestimation in the third leg.

When encountering additional weather in course C2, the number of overestimations increases from zero to four. This is substantiated by a significant effect of situation on the number of overestimations, $\chi^2(2) = 8$, $p < 0.05$. There is no significant effect of display on the number of overestimations, $\chi^2(2) = 4.6667$, $p = 0.097$. However, a clear trend is visible. Most overestimations took place when using the advisory display (3/7), followed by the constraint-based display (1/7). The difference between the baseline and the advisory display is particularly striking, as the advisory display contains all information of the baseline display, as well. Still, in this situation, having access to more information through the advisory display led to worse decisions. All three overestimations took place when the pilots chose the less optimal direction of circumnavigation, i.e., made a non-optimal trajectory decision.

The second outcome occurs when pilots underestimate their capabilities. In this case, they judge their fuel as insufficient to complete the next leg, even though the remaining fuel would actually be sufficient to complete the shortest one-turn solution to the next target. No significant effects can be observed. Pilots underestimated their capabilities twice with the baseline display. It is important to note that underestimations are not necessarily wrong, but can also be a sign of caution. It is possible that the pilots realised that the fuel is theoretically sufficient to complete the leg, but their previous experience taught them that they typically require a certain additional amount of fuel per leg.

The third outcome occurs when the prediction is accurate at the beginning and true in the end, i.e., the pilots neither overestimated nor underestimated their capabilities. For course C1, there is a significant effect of leg number on the number of true predictions, $\chi^2(2) = 6$, $p < 0.05$. Three wrong estimations took place in leg 3: two overestimations (one with the baseline, one with the constraint-based display) and one underestimation (with the baseline display).

For course C2, there is a significant effect of leg number ($\chi^2(2) = 10$, $p < 0.01$). Five wrong estimations occurred during leg 2, when the additional weather area appeared. One of these was an underestimation with the baseline display. The remaining wrong estimations were overestimations: three with the advisory, one with the constraint-based display. While the test statistic is not significant, $H(2) = 6$, $p = 0.050$, there seems to be a trend of a larger number of wrong estimations when using the advisory display in this situation.

The pilots were also asked to predict their capability of completing the remainder of the course, at the beginning of leg 1 and leg 2. Considering full-course overestimations in course C1, there is a significant effect of the employed display ($\chi^2(2) = 6$, $p < 0.05$). The baseline display caused four, the constraint-based display five overestimations, while the advisory display caused none. In situations without additional bad weather, the advisory display seemed to be very helpful in determining the remaining course capabilities.

Naturally, there is a significant number of overestimations in leg 1 of course C2, compared to the second leg, $\chi^2(1) = 14$, $p < 0.001$. These are caused by the not yet visible additional weather areas in leg 2. In this case, the pilots were acting on incomplete information and were unable to provide accurate estimations. Across both legs, the effect of display is significant, too ($\chi^2(2) = 8.4$, $p < 0.05$). However, analysing only leg 2 reveals no significant effects of display, $H(2) = 4.6667$, $p = 0.097$. The advisory display caused three, the constraint-based display two overestimations. The baseline display caused none.

Full-course underestimations happened rarely: twice with the advisory, twice with the constraint-based display. No significant effects are observed.

When pilots neither overestimated nor underestimated their full course capabilities, their prediction was true. For course C1, no significant effects are observed, $p > 0.05$ for every variable. For course C2, both display ($\chi^2(2) = 8.4$, $p < 0.05$) and leg number ($\chi^2(1) = 14$, $p < 0.001$) have a significant effect. As previously explained, the small number of correct estimations in leg 1 is explained by the additional weather area that would appear at leg 2. At leg 2, the baseline display caused zero wrong predictions, while the advisory display caused four and the constraint-based display caused two. The differences are not significant, $H(2) = 4.6667$, $p = 0.097$.

Fig. 12 shows the questionnaire results covering the fuel reserve estimation support of the displays. The advisory and constraint-based displays are rated significantly higher than the baseline dis-

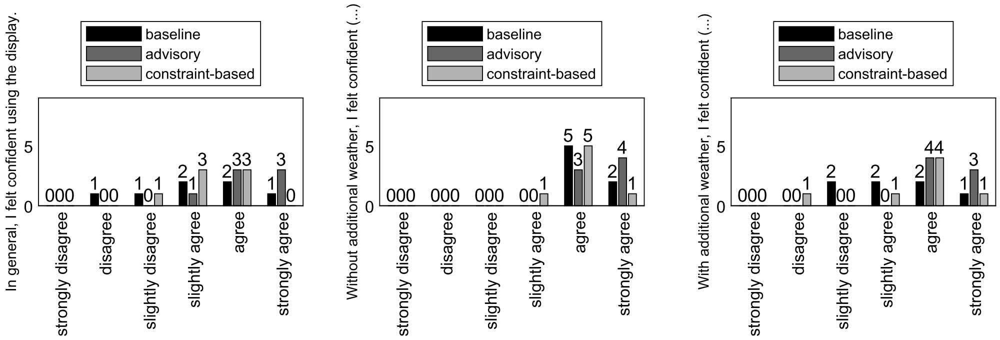

Fig. 13. I felt confident using the display. In general (left); without additional weather (middle); with additional weather (right).

play (H(2) = 14.1411, $p < 0.001$). Again, this difference is not visible in the objective experiment metrics, where each display contributed equally to some false predictions. In this case, having the additional information of the advisory or constraint-based display might have increased the pilots' confidence in their predictions, or decreased the required workload, without influencing the prediction accuracy.

6.3. Workload and situation awareness

The NASA-TLX workload rating is shown in Fig. 11 on the right-hand side. No significant effects are observed. There seems to be a slight trend of a lower workload rating with the advisory display in 2-turn courses, which would be in line with the hypothesis. Considering all conditions, however, the subjective workload seems largely independent from both situation and display. The situation awareness ratings are shown in the same Fig. 11. No significant effects are observed. There is a small trend of an increased situation awareness rating in 2-turn situations with the advisory display, mirroring workload ratings.

6.4. Pilot preference

Pilot preference was determined through questionnaires, as shown in Fig. 13. Both display ($\chi^2(2) = 11.9433$, $p < 0.005$) and situation ($\chi^2(2) = 6.6141$, $p < 0.05$) significantly impact the approval rating. In general, pilot preference was largest for the advisory display. In nominal situations, all displays were preferred equally, no significant effect can be observed (H(2) = 3.5, p = 0.1737). However, in cases with additional bad weather, the results diverge. Analysing the displays separately, encountering additional bad weather reveals a trend of decreasing preference rating of the baseline display (H(2) = 4.4842, p = 0.1062). Both the constraint-based display (H(2) = 3.0054, p = 0.2225) and the advisory display (H(2) = 0.5832, p = 0.7471) retain their high ratings.

This result is surprising, as the advisory display is in fact the least helpful in these situations, purely based on the information it provides. The trajectory advice it gives in these situations is obviously wrong (and the pilots were aware of this) and the suggested direction leads to a larger circumnavigation manoeuvre. Nonetheless, particularly in these situations, pilots preferred the “wrong advice” of the advisory display both over the minimal but correct information of the baseline display and over the partly wrong, but still usable information of the constraint-based display. This result is in stark contrast to the hypothesis as well as to the other dependent measures.

7. Discussion

The most important finding of this experiment is the influence of the employed automation system on the decision-making of the participating pilots. Results show that employing the advisory display significantly and negatively impacts the pilots' decision-making process in both off-nominal situations: a possible two-turn solution and an additional appearing obstacle. Critically, most pilots were unaware of their worse decisions, in particular during situations with possible two-turn solutions. A few pilots commented mid-run on their decision (e.g., *“I think I should have gone the other way.”*, *“Why did I fly this way around? I think through the middle would have been faster.”*) and some more pilots made references to “wrong suggestions” in the final questionnaire.

When encountering additional bad weather, the advice was always obviously wrong, as it crossed through an additional weather area. However, even in these situations, pilots were inclined to follow the direction of the suggested trajectory. This might be explained through a “priming” effect of the previously correct and still visible trajectory suggestion. To change their opinion about which direction to fly, the pilots were required to abandon the convenience of control Strategy (6) (and the large shortcut through the DL) and use Strategies (1) or (2). Being used to a nicely presented solution through the advisory display, they might have been inclined to utilise Strategy (2), choosing the first viable solution, instead of utilising the more mentally demanding Strategy (1), choosing the optimal solution. Being aware of the incorrectness of the large shortcut through the DL, the pilots might have been primed to select a control strategy that still provides the largest shortcut possible, without violating the safety requirements. By utilising this shortcut, they (possibly subconsciously) sacrificed trajectory efficiency for a quicker trajectory determination control strategy.

In this case, the first viable solution seems to be influenced by the still visible, albeit wrong, advisory suggestion. The solution that is closely related to this wrong suggestion is the trajectory that follows the same direction and just incorporates an additional track around the newly appeared obstacle. Half of the pilots implemented this suboptimal trajectory.

The pilots preferred the direct trajectory suggestion of the advisory display, in particular in situations with additional bad weather, where this advice was clearly wrong. At first glance, this seems to be a contradiction: in situations where the advice was clearly wrong, pilots preferred the display even more. Conversely, the baseline and constraint-based display were rated less favourably, even though the information of the baseline display was not influenced by the additional weather area and most of the constraint-based information was still usable. Clearly, the differences in approval rating are not based on the usability of the provided information alone.

In general, the way in which the question is posed in the questionnaire might have caused a difference. Even though there was no (baseline) or little (constraint-based) difference in support between situations with and without additional bad weather, pilots might have felt inclined to rate the displays worse in situations with additional bad weather. This might have been the case because this situation theoretically presented a complication in determining the next trajectory. In case of the advisory display, this negative effect might have been counteracted by the natural "convenientness" of a nicely presented trajectory suggestion, even if the suggestion was wrong. As previously discussed, this wrong suggestion could have acted as a primer to determine the "next best" solution quickly and without the requirement to completely change the control strategy and put more cognitive effort into it.

The constraint-based display did not cause significant differences in dependent measures, when compared to the baseline display. In terms of decision-making, there were only two instances of "wrong" trajectory decisions. Pilots commented on a variety of advantages and disadvantages of the constraint-based display. Some examples include:

- *"The display is hard to learn and use, caused by the novelty of the data type and representation"*
- *"The data is appreciated, but using it correctly causes a lot of workload"*
- *"It gives a lot of information, but you must calculate and do too much."*

The pilots appreciated the goal of the display and commented on the theoretical usefulness of the information. However, using the display correctly was hard to learn and caused high workload. Consequently, usage of the display differed between pilots. Some pilots used the display extensively, setting "test waypoints" to the right and left of obstacles to support their decision-making. Conversely, some pilots did not set intermediate waypoints at all and only utilised the information about additional travel distance signified by the ellipses. In addition, some pilots commented on the usefulness of the "future leg constraints", whereas other pilots did not use this information. In conclusion, it seems like this display concept may have potential, but it requires further improvement in terms of data representation, caused workload, and training.

This experiment clearly shows the disadvantages of subjective ratings of situation awareness and system preference. When analysing situation awareness, it is impossible for the participants to judge the amount of information they did not perceive or understand. They might have the subjective sense of perceiving and understanding all necessary information, as they are, obviously, not aware of all the information they did not perceive. This might be a reason for the relatively high values of SART ratings for the advisory display: the pilots are presented with an easy-to-follow suggestion, which causes a sense of solving the situation quickly and efficiently. Consequently, the subjectively reported SART ratings are high. However, the pilots are, at this moment, not aware of the existence of a more efficient trajectory. While this might have been caught by more objective measures of situation awareness, the subjective scale in this experiment has no way of incorporating this "unknown unknown".

Subjective measures of system value or acceptance can be misleading, as seen in this experiment. The advisory display led to the worst decisions but was most favoured by the pilots. It is unclear whether the pilots preferred the advisory display despite the worse decisions, or because they were unaware of the negative impact on their decision-making. It is therefore important to consider the state of information the pilots have while judging their system preference: does it incorporate the pilots' performance, or is it solely based on convenience? This is exceptionally important when designing new displays for actual operation: the convenience or ease-of-use of an automation system clearly does not correlate with the quality of its support and the resulting system performance and safety.

This is not a new result. The possible downsides of automation have been discussed extensively [1,26]. However, it is important to analyse and discuss these general findings while taking into consideration the actual system design and characteristics. In case of this experiment, it is important to consider the typical human-machine interface of helicopters and how its characteristics could influence the effect of automation systems.

In this experiment and in many helicopters that are used privately and commercially, pilots are required to fly "hands-on" the vast majority of a mission. As such, they have little to no capacity to use intricate touch-screen functions or other mechanisms that would require them to let go of one of the control inceptors. Because they are required to manually control the vehicle, it is conceivable that this human-machine system is increasingly susceptible to ironies of automation that discuss trust in automation and automation reliability, as described by Bainbridge [1]. As the manual workload of controlling the (typically unstable) helicopter is already high, any form of automation function that reduces the requirements on the pilots' mental capacities might be strongly appreciated.

This is particularly true for the advisory display of this experiment: it completely solves the trajectory determination task for the pilots, greatly reducing the strain on their cognitive resources. Therefore, pilots might be inclined to accept the suggested solution even if it could theoretically be (or actually is) wrong. The results of this experiment indicate that even when the suggested solution is clearly wrong, it can still be appreciated as a "starting point" for the determination of the next trajectory. Conversely, the constraint-based display places more requirements on the cognitive resources of the pilots. In these cases, even if the provided information is theoretically useful, the pilots might rather not use it and rely on the easier usage of baseline display information. This effect might have been exacerbated by the relatively short training time for both the advisory and constraint-based display functionalities. Pilots are naturally familiar with using a baseline navigation display, but all additional functionalities were novel, and their appropriate use needed to be learned. This effect possibly impacted the pilots' performance with both experimental displays.

Lastly, the requirements of continuously controlling the helicopter might impair the capacity of pilots to analyse the response of the used automation systems and to examine their behaviour in search for inconsistencies. It is conceivable and supported by the results of this experiment that pilots differentiate between different degrees of accepted automation failure. 6/8 pilots accepted less efficient trajectories when encountering possible two-turn solutions, resulting in worse mission efficiency. However, only 4/8 pilots accepted flawed suggestions when encountering additional weather, the rest of the pilots invested more cognitive resources to modify their decision. Lastly, none of the pilots followed a flawed suggestion into a bad weather area: analysing and correcting the automation suggestion always took preference over impairing safety.

8. Conclusion

This paper experimentally compared a baseline, advisory, and constraint-based helicopter navigation display. The eight participating helicopter pilots preferred the advisory display, even in situations where its advice was wrong. The advisory display caused the most suboptimal trajectory decisions, in particular in situations that afforded a more efficient trajectory without introducing additional navigational hazards. The complex information that the constraint-based display provided was appreciated by the pilots. However, this did not result in improved decision-making and the pilots commented on the difficulty of learning and utilising the display. The negative impact of the advisory display was clearly visible and warrants intense scrutiny for future system designers to avoid these negative influences of future automation systems in actual operation. Pilot remarks have highlighted the potential of the constraint-based display of this experiment to better inform pilots decisions. However, its data representation and the required high workload to correctly use it barred it from being more useful than the baseline display.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

We would like to thank all experiment participants for investing their time to enable this research. This project has received funding from the European Union's Horizon 2020 research and innovation program under the Marie Skłodowska-Curie grant agreement N° 721920.

References

[1] L. Bainbridge, Ironies of automation, Automatica 19 (6) (1983) 775–779, https://doi.org/10.1016/0005-1098(83)90046-8.

[2] J. Kaletka, H. Kurscheid, U. Butter, FHS, the new research helicopter: ready for service, Aerosp. Sci. Technol. 9 (5) (2005) 456–467, https://doi.org/10.1016/j.ast.2005.02.003.

[3] Anonymous, Crashed during approach, Boeing 737-800, near Amsterdam Schiphol Airport, 25 February 2009, Tech. rep., The Dutch Safety Board, 2010.

[4] K.J. Vicente, J. Rasmussen, Ecological interface design: theoretical foundations, IEEE Trans. Syst. Man Cybern. 22 (4) (1992) 589–606, https://doi.org/10.1109/21.156574.

[5] J. Comans, C. Borst, M.M. van Paassen, M. Mulder, Risk perception in ecological information systems, in: M.A. Vidulich, P.S. Tsang, J.M. Flach (Eds.), Advances in Aviation Psychology, 1st edition, Ashgate Publishing Ltd, Burlington, 2014, pp. 121–138, Ch. 8.

[6] M.P. Jenkins, C. Hogan, R. Kilgore, Ecological display symbology to support pilot situational awareness during shipboard operations, in: Proceedings of the 2015 IEEE International Multi-Disciplinary Conference on Cognitive Methods in Situation Awareness and Decision, Institute of Electrical and Electronics Engineers, 2015, pp. 213–219.

[7] D. Friesen, C. Borst, M.D. Pavel, O. Stroosma, P. Masarati, M. Mulder, Design and evaluation of a constraint-based head-up display for helicopter obstacle avoidance, J. Aerosp. Inform. Syst. 18 (3) (2021) 80–101, https://doi.org/10.2514/1.i010878.

[8] R. Klomp, C. Borst, R. Van Paassen, M. Mulder, Expertise level, control strategies, and robustness in future air traffic control decision aiding, IEEE Trans. Human-Mach. Syst. 46 (2) (2016) 255–266, https://doi.org/10.1109/THMS.2015.2417535.

[9] M. Mulder, R. Winterberg, M.M. van Paassen, M. Mulder, Direct manipulation interfaces for in-flight four-dimensional navigation planning, Int. J. Aviat. Psychol. 20 (3) (2010) 249–268, https://doi.org/10.1080/10508414.2010.487010.

[10] T. Ebel, Technical note: obstacle databases for instrument flight: a risk analysis, J. Aerosp. Inform. Syst. 16 (10) (2019) 437–440, https://doi.org/10.2514/1.I010737.

[11] A. Murrieta-Mendoza, R.M. Botez, Methodology for vertical-navigation flight-trajectory cost calculation using a performance database, J. Aerosp. Inform. Syst. 12 (8) (2015) 519–532, https://doi.org/10.2514/1.I010347.

[12] E. Greenwood, R. Rau, A maneuvering flight noise model for helicopter mission planning, J. Am. Helicopter Soc. 65 (2) (2020) 1–10, https://doi.org/10.4050/JAHS.65.022007.

[13] Y. Lim, A. Gardi, R. Sabatini, S. Ramasamy, T. Kistan, N. Ezer, J. Vince, R. Bolia, Avionics human-machine interfaces and interactions for manned and unmanned aircraft, Progr. Aerosp. Sci. 102 (2018) 1–46, https://doi.org/10.1016/J.PAEROSCI.2018.05.002.

[14] E. Guillanton, S. Germanetti, System architecture for new avionics on eurocopter fleet based on IMA supporting civil and military missions, in: Proceedings of the 37th European Rotorcraft Forum, Cascina Costa, Italy, 2011.

[15] S. Haisch, S. Hess, S. Jank Kreitmair-Steck, Pilot assistance for rotorcraft, in: Proceedings of the 35th European Rotorcraft Forum, Hamburg, Germany, 2009.

[16] M.D. Takahashi, M.S. Whalley, H. Mansur, C.R. Ott, J.S. Minor, Z.G. Morford, C.L. Goerzen, G.J. Schulein, L.C. Ott, M.J.S. Minor, M.Z.G. Morford, Autonomous rotorcraft flight control with multilevel pilot interaction in hover and forward flight, J. Am. Helicopter Soc. 62 (3) (2017) 1–13, https://doi.org/10.4050/JAHS.62.032009.

[17] S.M. De Bernardi, M. Ferroni, Skyflight mobile: a service to enhance the Leonardo flying experience, in: Proceedings of the 44th European Rotorcraft Forum, Delft, the Netherlands, 2018.

[18] C. Roos, A. de Reus, Effects of on board tablet use during helicopter operations, in: Proceedings of the 41st European Rotorcraft Forum, Munich, Germany, 2015.

[19] J. Rasmussen, Skills, rules, and knowledge; signals, signs, and symbols, and other distinctions in human performance models, IEEE Trans. Syst. Man Cybern. SMC-13 (3) (1983) 257–266, https://doi.org/10.1109/TSMC.1983.6313160.

[20] O. Stroosma, M.M. van Paassen, M. Mulder, Using the SIMONA research simulator for human-machine interaction research, in: Proceedings of the AIAA Modeling and Simulation Technologies Conference and Exhibit, AIAA, Austin, Texas, United States of America, 2003.

[21] I. Miletović, D.M. Pool, O. Stroosma, M.D. Pavel, M. Wentink, M. Mulder, The use of pilot ratings in rotorcraft flight simulation fidelity assessment, in: Proceedings of the American Helicopter Society 73rd Forum, Fort Worth, Texas, United States of America, 2017.

[22] S.G. Hart, L.E. Staveland, Development of NASA-TLX (Task Load Index): results of empirical and theoretical research, Adv. Psychol. 52 (C) (1988) 139–183, https://doi.org/10.1016/S0166-4115(08)62386-9.

[23] R.M. Taylor, Situational awareness rating technique (SART): the development of a tool for aircrew systems design, in: Proceedings of the AGARD AMP Symposium on Situational Awareness in Aerospace Operations, CP478, 1989.

[24] R.G.J. Miller, Simultaneous Statistical Inference, Springer Series in Statistics, Springer, New York, 2012.

[25] C. Borst, M. Mulder, M.M. Van Paassen, Design and simulator evaluation of an ecological synthetic vision display, J. Guid. Control Dyn. 33 (5) (2010) 1577–1591, https://doi.org/10.2514/1.47832.

[26] B. Strauch, Ironies of automation: still unresolved after all these years, IEEE Trans. Human-Mach. Syst. 48 (5) (2018) 419–433, https://doi.org/10.1109/THMS.2017.2732506.