

Mixed-Initiative Control of a Roadable Air Vehicle for Non-Pilots

Michael C. Dorneich¹, Emmanuel Letsu-Dake
Honeywell Laboratories

and

Sanjiv Singh, Sebastian Scherer, Lyle Chamberlain,
Marcel Bergerman
Carnegie Mellon University

This work developed and evaluated a human-machine interface for the control of a roadable air vehicle (RAV), capable of surface driving, vertical takeoff, sustained flight, and landing. Military applications seek to combine the benefits of ground and air vehicles to maximize flexibility of movement but require that the operator have minimal pilot training. This makes the operator vulnerable to automation complexity issues; however, the operator will *expect* to be able to interact extensively and control the vehicle during flight. A mixed-initiative control approach mitigates these vulnerabilities by integrating the operator into many complex control domains in the way that they often expect—*flexibly in charge, aware, but not required to issue every command*. Intrinsic safety aspects were evaluated by comparing performance, decision making, precision, and workload for three RAV control paradigms: human-only, fully automated, and mixed-initiative control. The results suggest that the mixed-initiative paradigm leverages the benefits of human and automated control while also avoiding the drawbacks associated with each.

Keywords: mixed-initiative control, roadable air vehicle, human-machine interface

Introduction

There is a long history of attempts to combine the advantages of a ground vehicle with those of an air vehicle, in order to provide more efficient and flexible mobility. A *roadable air vehicle* (RAV) is a road vehicle that has the capability to drive, take off, fly, and land. The goal of an RAV is to provide a capable hybrid vehicle with flexible aerial mobility that can be manually driven on the ground. With such a vehicle, transportation will no longer be restricted to trafficable terrain that tends to make movement predictable. Moore (2006) describes the requirements for a “Third Wave of Aeronautics” as one of on-demand mobility. Basic requirements include computer-based guidance and intelligent crew decision support systems to enable operators to make decisions best suited to their needs and capabilities. In addition, an air vehicle would need to provide its own

¹ Now at Iowa State University.

surveillance to enable safe flight in the presence of other aircraft and environmental conditions (Moore, 2006).

Striking a compromise when merging automobile and aircraft design is difficult, as made clear by earlier efforts (Genta, Morello, Cavallino, & Filtri, 2014). The resulting vehicle is likely to be a poor road vehicle and a poor aircraft (Jump et al., 2011). In the civilian domain, the spectrum of RAV design runs from a vehicle that is primarily an airplane that can also drive on the ground (Terrefugia, 2012) to a vehicle that is designed to drive and can also fly (e.g., PAL-V, 2012). Commercial RAV applications, from the Taylor Aerocar prototype of the 1950s to the current Terrafugia Transition, require operators to be licensed pilots. For example, the Transition requires a light sport plane pilot license. While this has fewer requirements than a private pilot's license, use is limited to personal (non-commercial) applications and only in daylight visual flight conditions (Weissler, Mason, & Woock, 2012). Commercial applications, such as an air taxi, are further constrained, because system administrators and professional pilots would be required (Moore, 2006).

In military domains, an RAV needs to drive, fly (or some combination of the two for a tactically useful distance), take off, and land in difficult locations. A key requirement of an RAV is the safety of passengers during takeoff, flight, landing, and in emergencies (DARPA, 2010). The military is developing an RAV capability that is designed to be operated by a non-pilot, who must be able to command and control the vehicle safely within a range of modes that vary from direct, real-time to fully autonomous operator steering commands (DARPA, 2010). Unlike unmanned systems operated remotely, human operators in an RAV have the best and most immediate information about the weather environment, obstacles, traffic, and airborne resources. They are best-positioned to make decisions affecting the flight of the vehicle. Overall, the goal is to combine the advantages of automation, surveillance, sensor systems, and human situational awareness to minimize workload bottlenecks, simplify complex tasks, and increase task performance, all while maintaining robust safety in the face of failures and unexpected events (Moore, 2006).

The goal of the work described here is to develop a human-machine interface (HMI) for a non-pilot operator of an RAV. The requirement that the operator be a non-pilot greatly increases the requirements for both the automation and the human-automation interaction. Even though the operator is not required to have pilot training, we expect that some level of training beyond operating a conventional road vehicle will be required to operate the RAV. Apart from basic human-automation interaction, this assumed level of training will also account for the handling of failure mode conditions.

Critical Technologies in an RAV System

RAV control for the non-pilot operator will require three critical technologies: 1) *Perception*: sensor support to maintain the RAV operator's situational awareness of the terrain and airspace within which the vehicle operates; 2) *Planning*: automated planning technologies to support route planning, approach planning, landing zone (LZ) selection, and collision avoidance; and 3) *Human-Machine Interface (HMI)*: an intuitive control paradigm to allow non-pilot operators to exercise safe control within the realm of their capabilities. Considered together, technologies supporting the RAV operator's situational awareness, collision avoidance, LZ selection and validation, and intuitive control can provide the necessary safety and efficiency for the targeted missions. Careful design of the combination of human and machine capabilities will additionally produce the robustness necessary for military operation. Fig. 1 illustrates the three critical components of the RAV system: Perception, Planning, and the HMI.

The Perception system utilizes a sensor suite composed of a laser scanner, radar, cameras, and a priori data (e.g., maps, building locations) in order to keep the vehicle safe from collisions by identifying feasible landing sites free of obstacles and providing situational awareness for the

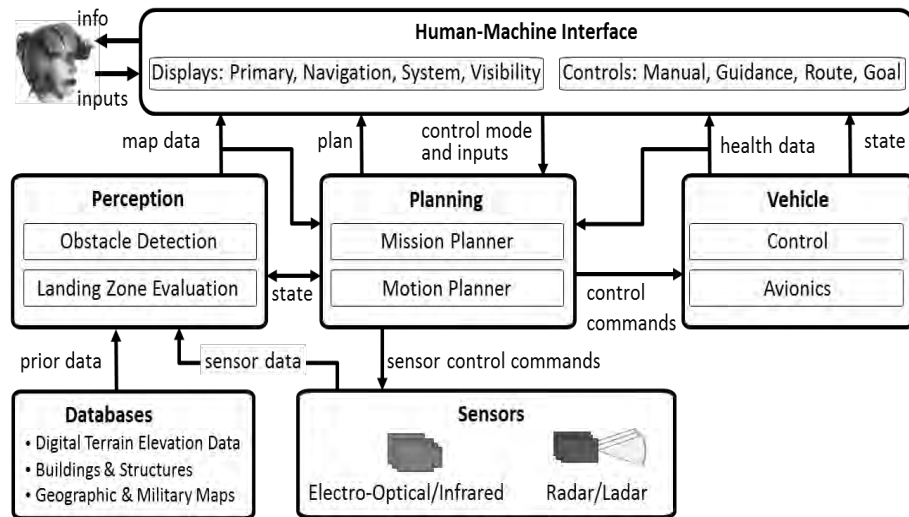


Figure 1. Functional architecture of the RAV autonomy system. This paper focuses on the HMI subsystem; detailed elements are described in the remainder of the paper.

vehicle's operators. Perception will build up a terrain map and highlight problem areas as an overlay on a map or on a camera view. The LZ must be continuously monitored for moving objects, such as people and vehicles.

The Planning system can route an RAV safely in a variety of modes. When in flight, the Motion Planner can provide velocity controls to keep the vehicle from executing unsafe maneuvers. The Mission Planner produces routes in response to human specification of destinations; the generated route will adhere to known constraints such as no-fly zones, altitude constraints, and viable approach paths to landing sites. When obstacles are detected by the Perception module, the planning system can react quickly to avoid a collision.

The HMI allows the operator to direct the vehicle through a set of layered controls that can vary from manual (direct control of the aircraft) to fully autonomous. The HMI displays help an operator maintain a situation awareness of the vehicle state, the environment, and the mission. The HMI will be discussed in detail in the remainder of the paper.

Unmanned air vehicles (UAVs) are commonplace today, but most take off and land on prepared surfaces (e.g., runways and helipads). The difficulty of automating takeoff and landing on unprepared surfaces and in proximity to buildings, trees, and wires has kept autonomous rotorcraft behind their fixed-wing counterparts. RAVs need exactly such capabilities: to be useful in performing tasks like resupply or medical evacuation, they must be able to quickly take off and land at unprepared sites. This need cannot be met simply by either execution of pre-planned trajectories or by incorporation of a fully manual mode. Small errors in maps and changing environmental conditions can foil simple autonomy. Additionally, there are critical functions that the vehicle cannot perform without human input. For instance, while sensors can detect the shapes of obstacles in an LZ, they cannot determine the suitability of the surface for landing (flat rock vs. mud, for instance). It will therefore be up to the operator to make the final landing feasibility decision.

Human-Machine Interface Challenges

Safety is a key priority for all aviation and a particularly important issue for an RAV, because: (1) operators are not expected to have pilot training; (2) operators will not have the experience to deal with unexpected cases; and (3) autonomous aerial vehicles (manned or unmanned) are two to three orders of magnitude less reliable than commercial aircraft. With respect to (2), the types of use cases RAV operators may have to contend with include, among others:

- Mechanical vehicle failure during flight.
- Main sensor failure (e.g., leading to impossibility of validating an LZ).
- Sensor error (e.g., leading to the LZ being incorrectly marked as safe/unsafe, obstacles en route not being detected, etc.).
- Map error (e.g., leading to the LZ being incorrectly marked as landable).
- Objects present in selected the LZ (e.g., vehicles, animals, personnel, cargo).
- Other vehicles (blue or red forces) flying on collision course.
- Operator incapacitation.

The safe operation of the RAV will require human control in both normal and off-normal operations. As a human-rated flying vehicle, it must meet safety and reliability goals that are not achievable by automation alone. Our approach to mitigating these vulnerabilities is to design with the role of the operator as a goal setter and a flight manager, not as a pilot. It is the joint human-automation control of the vehicle that will enable the safe operation and full realization of operational capabilities. The goal of the HMI is to support safe and reliable operations by supporting human involvement in the control of the vehicle at whatever level necessary for optimal achievement of mission objectives.

Additionally, the driving requirements of the vehicle introduce constraints affecting the operator's visibility during flying. The operator should be able to view and assess the precise landing point in relation to terrain and hazards, but the physical layout of the vehicle will occlude most of the traditional out-of-the-window (OTW) views expected in rotorcraft. Thus, there is a degraded situational awareness due to the loss of visual cues. We address this problem by adding the capability for the operator to see "through" the vehicle via sensors mounted outside of it.

This paper describes work to develop and evaluate a human-machine interface display and control paradigm that allows a non-pilot to control an RAV in a safe and efficient manner while fulfilling a wide variety of mission goals. The human-machine interface implements a mixed-initiative control philosophy. The next section describes the rationale for our technical approach, followed by a brief description of the system. Subsequent sections detail an evaluation that assessed intrinsic safety aspects of the system by comparing performance and workload for three control paradigms for the RAV: direct human-only control, automation-only control, and a mixed-initiative human-automation control. As the most dangerous phases of flight are takeoff and landing, this work first focused on how to design an HMI that manages the complex interplay between the perception capabilities, the planning and control automation, and the HMI to support the non-pilot operator during landing. This is followed by the results section, which highlights and discusses key findings of the evaluation. The conclusion section captures highlights of the study and future research implications.

Background

The Challenges of Manual or Automated Control

There is a continuum of possible control methodologies that could be developed for the RAV. At one end of the spectrum, the vehicle would be under full human control, but the profile of a non-pilot RAV operator precludes this option. The operator has no flight training and, most likely, a

minimal understanding of the physics of flight. Additionally, the operator will have varying levels of attention to devote to vehicle operation (given the imperatives of the mission). The lack of flight training also makes the RAV operator vulnerable to other automation complexity issues in general and aviation automation issues in particular. These include:

- Breakdowns in automation mode awareness due to interaction complexity.
- Integration of information from multiple sources.
- Vulnerabilities due to lack of operational knowledge (for example, speed and energy management).

At the other end of the continuum, the vehicle could be fully automated. However, three reasons preclude this as a viable alternative: 1) certification requirements of a human-rated vehicle, 2) the sensor requirements for fully automated approach and landing, and 3) the inherent brittleness of automation design in complex domains.

Since the operator is not a pilot, the vehicle automation should have a measure of responsibility for its own safety, detecting and avoiding obstacles, and finding safe LZs. As a human-rated vehicle, however, it must meet stringent safety and reliability targets on the order of 10^{-9} probability of occurrence of each catastrophic failure condition per hour of flight (FAA, 1988). However, the most advanced unmanned flying vehicles have only achieved 10^{-4} per hour mishap rate (Barnard Microsystems Limited, 2011). The certification requirements alone would make a human-rated but not piloted vehicle, such as an RAV, prohibitively expensive.

RAVs need perception and planning capabilities to be useful in performing tasks like resupply or medical evacuation, and they must be able to quickly take off and land at unprepared sites. This need cannot be met simply by either automated execution of pre-planned trajectories or by incorporation of a fully manual mode.

Finally, research has indicated that highly automated systems suffer from brittleness (the system model does not account for all possible scenarios) (Smith, McCoy, & Layton, 1997). For decades, *supervisory control* (Sheridan, 1987) in complex and dynamic domains has remained such a challenge that many taxonomies of human interaction in control ignore operator instruction as a methodology (Parasuraman, Sheridan, & Wickens, 2000). Often, humans are seen as the primary backup when highly automated systems fail. But it is important to recognize the repercussions of designing a system that by necessity relies on the human only to take over automated functions in the case of automation failure. Reason (1990), for example, suggests that human operators are necessary primarily because designers neither foresee all operational situations nor provide automated measures for all contingencies. Humans, however, are ill-suited for passive monitoring (Parasuraman & Bowers, 1987). Another concern is the failure to appreciate the interaction between sub-task execution and management activities, such as mission planning, objective setting, and managing emergencies (Abbott, 1993; Alter et al., 1995). If one considers only performance of isolated functions and tasks in determining function allocation between humans and automation, many tasks may be automated that are relevant to mission planning and awareness, so users are left uninformed and uninformed, compromising their ability to perform higher-level mission planning tasks.

Mixed-Initiative Control

In contrast, *mixed-initiative control* methods show great promise for integrating the human into many complex control domains in the way the human often wants to be integrated—*flexibly in charge and aware but not required to issue every command* (Miller, Funk, Dorneich, & Whitlow, 2002; Riley, 1989). Based on this design philosophy, the HMI supports several RAV flying tasks, and in particular, cooperative tasks between the operator and the system. Cooperative tasks include awareness of the task allocation, RAV control mode, current operator state, and future capacities and availabilities of the operator. Some previous studies discuss aspects of these types of HMI, for example, shared and common work spaces that enable the building and maintenance

of a shared or mutual cognitive environment (Rognin, Salembier, & Zouinar, 2000; Pacaux-Lemoine & Debernard, 2000). Through mixed-initiative control, the tactical and strategic levels are allocated to the operator or the automation system according to their abilities or availabilities. This allows the limited mechanical and information processing capacities of the human operator to be compensated by means of cooperation with automation (Schmidt, 1990).

Mixed-initiative interaction was first defined in the computer-assisted instruction work of Carbonell (1970). Although a thorough review is beyond the scope of this paper, a short review of some definitions is relevant. Carbonell identified the needs for two-way interaction where both the human and the automation were required to contain knowledge and maintain understanding. Donaldson and Cohen (1997) defined elements associated with initiative: control, goals, and turn-taking. Allen (1999) referred to mixed initiative as a flexible interaction strategy where any agent may have the initiative to control the interaction and can dynamically adapt to best address the problem at hand.

Central to mixed-initiative systems is the negotiation between the human and the automation, resulting in the joint system taking advantage of each of their skills, capabilities, and knowledge to best adapt to contingencies in the environment during task execution (Landén, Heintz, & Doherty, 2012). The human and the automation share responsibility of mission safety and success (Goodrich, Olsen, Crandall, & Palmer, 2001). It is important to distinguish between responsibility and authority. Flemisch et al. (2012) propose that applied to human-machine systems, the authority of an actor can be defined by what the actor is or is not allowed to do, and is usually determined beforehand at design time. Responsibility indicates which actor is responsible for the outcome of a specific function or task is also usually determined during design. Research has shown that the allocation of partial functionality to automation actually changes the nature of the work for the operator, because the assignment and responsibility have been split (Billings, 1997; Miller & Parasuraman, 2007; Parasuraman et al., 2000; Parasuraman & Riley, 1997; Woods, 1996). The split between assignment and responsibility adds additional “induced” functions, such as monitoring, communicating, and coordinating (Lee, Kim, & Feigh, 2009).

The essence of mixed-initiative interaction is the fluid trading of multiple levels of control between the human and the automation. The result is increased flexibility, more robust joint human-automation system behavior, and improved human involvement. (Kortenkamp, Bonasso, Ryan, & Schreckenghost, 1997). The levels of autonomy are adjustable, and the mixed-initiative interaction is the process with which to move between levels. The non-pilot RAV operator requires a flexible, adjustable level of control to account for the range of situations that he or she will encounter in controlling the RAV (Fig. 2). The mixed-initiative interaction approach is appropriate for the RAV domain for several reasons: 1) both the operator and the vehicle are responsible for maintaining safety and mission success, 2) the system is designed to take advantage of the strengths, capabilities, and knowledge of the non-pilot human operator and automation, and 3) control is traded in a fluid way as best fits contingencies of the situation.

The premise is that the non-pilot RAV operator will not be controlling flight surfaces to achieve desired end states, which requires a second- or third-order integration of the control laws (Beringer, 2002). In fact, due to the non-pilot nature of the RAV operator, the HMI design is focused on enabling the operator to focus only on “zero-order” commandable end states: heading, altitude, and speed. The RAV operator should be able to exercise control along a continuum of time horizons from real-time tactical to longer-term strategic guidance. The operator control along the continuum aligns with the cognitive control levels described by Rasmussen and Vicente (1989). As automation control increase across the continuum, operator control progresses from skill-based (full-human) to rule-based (mixed-initiative) and knowledge-based (full-automation).

The operator’s interaction with the RAV is analogous to the tele-operation of an autonomous vehicle, but control takes place from within it. The operator benefits from better situational awareness due to direct sensory inputs and a wider field of regard, which improves real-time

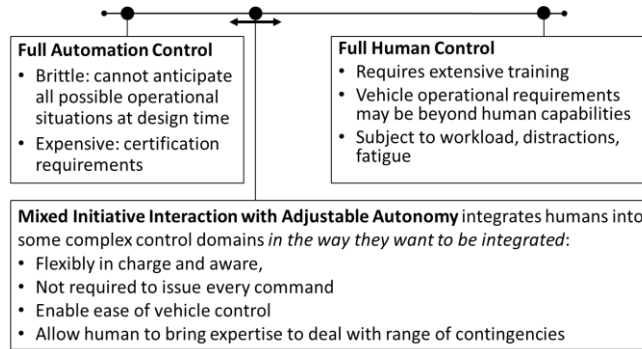


Figure 2. Continuum of possible control methodologies for an RAV.

decision-making. However, since the operator is subject to information overload, especially in high-tempo operations, the vehicle needs to be smart enough to understand what the operator is doing at any time and take safe actions, such as changing the levels of control if necessary. Here we use the “H-Metaphor” of horseback riding to explain our approach: “If you were riding a horse, you would be able to read your map and be confident that you would not hit any trees or run into people because horses instinctively avoid obstacles. If the horse is unsure about where to go, it may slow down, and seek a new obstacle-free path while trying to get the rider back into the loop” (Flemisch et al., 2003). The levels of safety and reliability demanded by the RAV will require the HMI design to leverage the flexibility and situational awareness of the human operator while simultaneously leveraging the reliability, predictability, and capability of automation. This approach also has the potential to maximize joint performance by improving decision making, reducing response time, and increasing operational tempo. The HMI will enable the safe operation of the RAV while minimizing human interaction, increasing situation awareness, and enabling improved control.

Human-Machine Interface Design

The human-machine interface design development was guided by the human-centered systems approach. Specifically, a top-down process identified design requirements from an analysis of mission descriptions. The team developed mission descriptions from operational requirements, with input from the Defense Advanced Research Projects Agency (DARPA), RAV vehicle design teams, and other subject matter experts (SMEs). Several mission descriptions were needed to cover the variations in operations for the RAV vehicle. These mission descriptions were captured as a set of Design Reference Missions (DRMs). The DRMs focused on normal operations and were designed in such a way as to exercise all the perception, planning, and HMI features and functions. The DRMs contained mission elements for potentially high workload situations during normal operations and during some failure situations (e.g., aborted landing, engine failure, etc.).

Functional analysis was used to define what the HMI must do to accomplish the DRMs. The functional analysis defined the functional behavior required for each operational event in the DRM. A list of top-level functions was developed and categorized.

Although flying in three dimensions (3D) is considerably more complex than driving in two dimensions (2D), the physical user interface must be essentially the same for both driving and flying. This enables the flying operation while maintaining the well-known driving controls

available in any road vehicle (steering wheel and pedals). To make flying intuitive to the operator, the number of controls is kept to a minimum. A notional layout of the human-machine interface (HMI) is shown in Fig. 3.

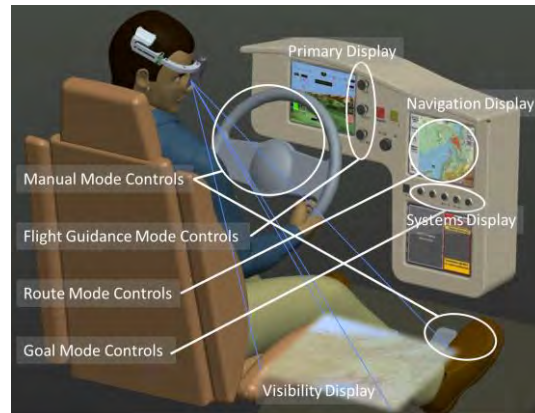


Figure 3. Layout of the RAV Human-Machine Interface.

In keeping with high-level operator functions, the *Primary Display* is in the operator's primary field of view and supports aviating and driving. There are two displays in the operator's secondary field of view. The *Navigation Display* supports the operator functions of navigation and the *Systems Display* supports systems management. The *Visibility Display* supports the operator's need to be able to see beyond the limits of the out-the-window view to identify feasible landing areas and may be on a head-mounted display or integrated onto one of the other existing displays.

Displays

Excessive information, much of it useful only to trained pilots, clutters displays and may confuse the non-pilot RAV operator. In our approach, only requisite information to support related tasks, flight control modes, and transitions between modes are presented.



Figure 4. The Primary Display (left) and Navigation Display (right).

Primary Display. The location of the Primary Display (PD) is directly in the operator’s field of view to help streamline instrument scanning. All primary situational information is consolidated on the PD to reduce operator scanning workload while maintaining the operator’s situation awareness. Analysis of the expected operator functions identified speed, altitude, heading, attitude, fuel, alerts, mode status, control limits, and controls feedback as the highest criticality information. The PD is illustrated in Fig. 4 (left).

Navigation Display. The Navigation Display (ND) supports the majority of the in-flight interaction with the operator. The operator has both tactical and strategic goals relating to navigation, and the ND must support both. To support operator navigation, the user interface supports path planning, vehicle position on a moving map, and layered displays (e.g., terrain, landmarks, LZs, traffic). The ND is illustrated in Fig. 4 (right). Functional analysis identified the following as information that should be persistently displayed to support operator situational awareness: location, terrain information, map range, heading, mission plan, and map features (e.g., terrain, roads). Information not located on the map is located in the dedicated status information area on the left of the display.

Systems Display. The Systems Display (SD) is intended to host information that operators can use to manage the various systems within the RAV. Examples include alert and notification displays, system diagnostic displays, and communications interfaces.

Visibility Display. The RAV will have significant blind spots that make it difficult for the operator to see outside with a wide enough field of view to make the final decision on takeoff and landing viability. While the perception sensor system will assist the operator to know when it is not safe to land, it is primarily the operator’s responsibility to decide when it is safe to land, and so the operator has final landing authority. To address this issue, a sensor suite consisting of laser scanner, radar, and cameras will be mounted around the vehicle (Fig. 5). A Near-to-Eye (NTE) display was developed to allow the operator to “see through the vehicle” for navigational landmarks when landing, thus eliminating the blind spots. The display is paired with a head tracker, so whichever way the operator turns his or her head, the display renders the outside scene in that direction.

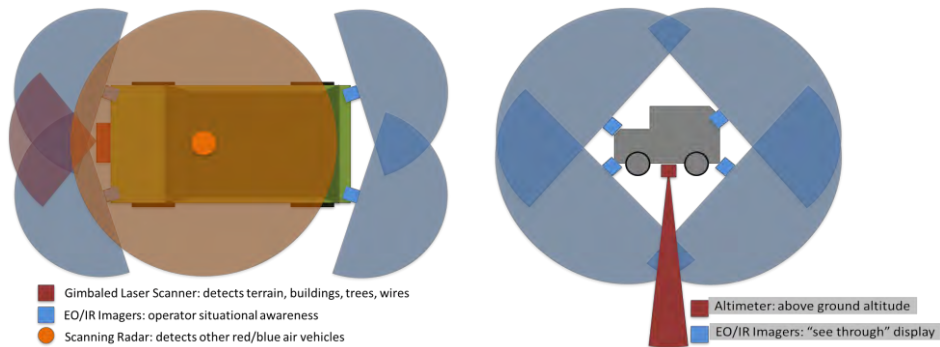


Figure 5. RAV sensor suite, including laser scanner, radar, and electro-optical and infrared (EO/IR) cameras.

Control Modes

The fundamental control philosophy of the RAV dictates that the non-pilot operator sets the goals for flight, and the automation is responsible for achieving those goals. Thus, in each control mode, the operator sets the desired speed, altitude, and heading. Higher order properties, for instance, climb rate, acceleration rate, and bank angle, are the responsibility of the automation, and therefore not under direct operator control.

The RAV HMI supports four distinct control modes (see Fig. 6) that offer flexibility of interaction based on factors such as situational awareness, workload, and mission demands:

- *Manual control mode*: This mode allows the user to directly move the vehicle left, right, up, down, faster, and slower.
- *Flight guidance control mode*: In this mode, the operator sets individual flight values of speed, heading, or altitude for execution by the automation.
- *Mission-based control mode*: This mode allows the creation, modification, or selection of flight routes and automation control of the vehicle based on the selected route.
- *Goal-based control mode*: This mode provides one-click shortcut commands, based on an intended operator goal understandable to both the operator and the automation.

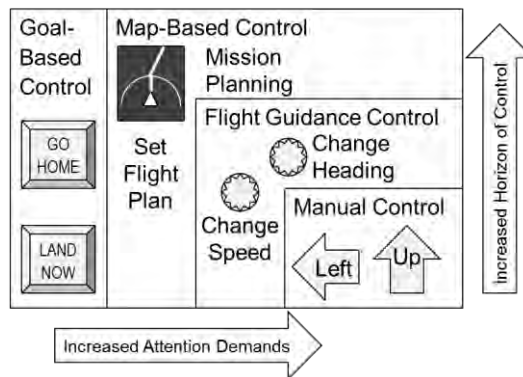


Figure 6. Layered flight control modes enable operators to tailor control to available resources and goals.

The modes vary in the amount of attention required by the operator (minimal in Mission-Based control, high in Manual Control) and the time horizon under which they are exerting control (long durations for Mission-Based and short for Manual Control). The Goal-Based control mode requires low attention but varies along the time horizon depending on the complexity of the commanded action.

Manual Control Mode. This control mode was designed to allow the user to move left, right, up, down, faster, and slower. The manual controls afford an experience similar to regular driving, using a steering wheel and pedals to control the flight parameters. Table 1 describes how to control the heading, altitude, and speed using the continuous manual mode controls.

Table 1. Discrete Manual Mode control options.

Control element	Flight Parameter	Action Description
Steering Wheel	Heading	Increase (+) in heading by turning the wheel to the right Decrease (-) in heading by turning the wheel to the left
Steering Wheel	Altitude	Increase (+) in altitude by pulling the wheel towards operator Decrease (-) in altitude by pushing the wheel away from operator
Foot Pedals	Speed	Increase (+) in speed by pressing the right foot pedal Decrease (-) in speed by pressing the left foot pedal

A small deflection in the control element will result in incremental parameter changes. Large deflections will result in accelerated changes in the flight parameter. Release of a control element zeroes the rate of change of the associated parameter.

Flight Guidance Control Mode. In this mode, the operator sets the individual flight values of speed, heading, or altitude. It is the automation's role to execute to the new goal and "fly" the vehicle to achieve the goals set by the operator. The objective of this mode is to allow the operator to give short-term guidance to the vehicle by changing its flight profile. It requires a more than a minimum level of attention but provides an opportunity for precise vehicle control.

Mission-Based Control Mode. This control mode allows creation, modification, or selection of flight routes. Once selected, the automation will fly the vehicle based on the loaded route. Route plans can be modified either before or during transit.

Building a Route. A Graphical Mission Planning (GMP) design was developed for the ND to allow operators to quickly plan a route when the destination is known a priori. Operators most likely will simply define a destination by selecting a point on a map via the touchscreen. The display highlights safe areas referencing a database of known safe landing areas. Alternately, the operator could select an area where the suitability for landing is, at planning time, unknown. Finally the user can refine his or her LZ selection during transit or upon arrival near the desired landing area. The operator can define four different types of waypoints: takeoff, landing, air points, and ground points. Air waypoints can be further defined with specific altitude and speed constraints. As in Manual Mode and Flight Guidance Mode, the operator has control over set points (speed, heading, altitude) to guide the automated flight control of the vehicle.

Approach and Landing. There are three sources of information to assess the viability of an LZ. The first source is a database of feasible landing sites. This database is used in the initial selection of the LZ during route creation or modification. The second source of information is the real-time information from the Perception subsystem, which scans for obstacles as the vehicle makes its final approach. Perception can identify obstacles, but it cannot identify the characteristics of the surface (e.g., hard sand, mud). Finally, the third source is the operator's visual inspection in order to confirm that the landing site is safe for landing.

In each approach profile, there is a point where the vehicle will not have enough energy to abort and must land. Up until that time, however, the operator can command an abort. A countdown timer is shown on the ND. The possibilities of perception, operator decisions, and resulting vehicle actions are shown in Table 2. Note that by default, the vehicle will not land unless commanded to do so by the operator.

Table 2. Planner and vehicle system actions based on perception and operator landing decisions.

Perception Results	Operator Response		
	Operator confirms landing	Operator triggers abort	No operator response
No obstacle detected (landing recommended)	Land at selected LZ	Default abort behavior (override landing)	Default abort behavior
Obstacle detected; (landing not recommended)	Land at selected LZ (override abort)	Default abort behavior	Default abort behavior

Goal-Based Control Mode. The Goal-Based Control Mode simplifies control of the vehicle by providing shortcut automated commands that allow the operator to quickly direct the vehicle while reducing workload, although at the expense of flexibility. Modes are activated by pressing labeled buttons located below the navigation display. The goals associated with each Goal-Based Control Mode are described in Table 3.

Table 3. Goals associated with the five goal-based modes.

Goal-Based Mode	Operator Goal
Take Off	Take off immediately, fly default profile until commanded with new operator input.
Loiter	Maintain safe and stabilized flight in a defined area.
Go To	Fly the RAV to a specified location.
Go Home	Fly the RAV to a predefined LZ.
Land Now	Quickly locate a safe LZ and land the RAV.

Evaluation

Objective

The objective was to evaluate the intrinsic safety aspects of the RAV HMI by comparing performance and workload under three different control paradigms for the RAV based on the continuum of possible control methodologies: human-only control, automated control, and a hybrid human-automation control.

Participants

There were six participants who each spent approximately four hours in the evaluation. Participants were recruited from the Carnegie Mellon Robotics Institute and had no pilot or flight training; they were therefore representative of the intended RAV operator. The median age was 30.5 years (range 26-39 years). Participants rated their video gaming experience on a four-point scale (None, Low, Average, or High). Although video gaming experience reported in literature is varied, the response anchors in this study were guided by Bailey, West, and Anderson (2010), in which video gaming experience was tied to gaming time per week (None: 0-2 hours per week; Low: 10 hours per week; Average: 20 hours per week; High: 40 hours per week). Video gaming experience of our six participants included none (one out of six), low (two out of six), and average (three out of six).

Independent Variables

The first independent variable is the *control paradigm* with three conditions: *Human*, *Automation*, *Human + Automation*. The three control paradigm conditions are defined in Table 4.

The second independent variable is the state of the landing zone (*LZ State*) with three conditions:

- *Clear LZ*, where the vehicle can safely land;
- *Obstructed LZ*, where there is an obstacle on the LZ that prevents safe landing; and
- *Unstable LZ*, where the surface of the LZ is not safe to land (e.g., a marsh).

The third independent variable is the *automation correctness*, with three conditions:

- *Automation Correct*, where automation correctly identifies an obstructed or clear LZ;
- *Automation Error of Commission*, where automation says LZ is obstructed when it is safe;
- *Automation Error of Omission*, where the automation says LZ is clear when it is obstructed.

Table 4. The three control paradigms used in the evaluation.

	Human (H)	Automation (A)	Human + Automation (H + A)
Operator	Controls vehicle in manual mode Visually assesses LZ (OTW and NTE) Makes land/abort decision When on top of the LZ, presses LAND NOW button to land	Does not perform any action related to flight or LZ assessment (i.e., operator is merely a “passenger”)	Assesses LZ both visually (OTW/NTE) and taking into account perception system return Makes final land/abort decision Able to override system’s landing decision
RAV Autonomy System	Disabled	Controls the vehicle in route-based or goal-based modes Scans LZ to make viability assessment Makes land/abort decision Lands the vehicle	Controls the vehicle in route-based or goal-based modes Scans LZ Provides operator assessment of LZ viability Lands the vehicle

As a result of these independent variables and the RAV design, there will be experimental conditions where human operators will be unable to override the automation, even if they knew it was in error. From a human factors perspective, this should never be the case for ideal human-automation interaction, since there should always be an option for the human operator to regain control authority. However, an underlying motivation of a RAV design for non-pilots is that future advancements in technology may be expected to ultimately assure the reliability of automation and minimize or eliminate imperfect automation. These conditions were included in this study merely for experimental purposes.

Hypothesis

It was expected that human-only control would be less precise and require higher workload from the operator but result in fewer decision errors when determining if it was safe to land or if an

abort was required. Under complete autonomous control, it was expected that the vehicle control would be more precise and the operator workload would be lower, but errors in the land/abort decisions would occur due to the inherent limitations of any sensor system. For instance, the current perception system can detect an obstacle on the LZ and inform the system and operator to abort. However, no automated system can be expected to be perfect. Furthermore, the perception system cannot detect an unstable LZ surface, so only the human operator can detect the unstable surface and command the vehicle to abort. Finally, it was hypothesized that a hybrid human-automated control would result in lower operator workload, higher control precision, and fewer decision errors. The hypothesis of the experiment is summarized in Table 5.

Table 5. The hypothesis of the protocol test.

Control Paradigm	Workload	Precision	Decision Making	Efficiency
(H) Human	Poor	Poor	Good	Poor
(A) Automation	Good	Good	Poor	Good
(H+A) Human + Automation	Good	Good	Good	Good

Experimental Design

This was a 3 (*Control Paradigm*) x 3 (*LZ State*) x 3 (*Automation Correctness*) within-subjects, repeated measures experimental design. There were 16 trials in each of the three control paradigm conditions (*Human*, *Automation*, *Human + Automation*). In each control paradigm, eight trials had a *Clear LZ*, four trials had an *Obstructed LZ*, and four trials had an *Unstable LZ*.

For this evaluation, we also wanted to study the ability of the human operator to compensate for imperfect automation, so we introduced a third independent variable of automation correctness, which included automation errors of commission and errors of omission. Thus, the 16 trials were divided into the five possible combinations of automation paradigm, LZ State, and automation correctness condition (Table 6). Note that since there was no perception feedback on LZ stability, there was no situation where automation could give feedback on an unstable LZ.

Table 6. Sixteen trials per control paradigm condition.

Landing Zone State	No Obstacle Detected	Obstacle Detected
(C) Clear LZ	7	1 (error of commission)
(O) Obstructed LZ	1 (error of omission)	3
(U) Unstable LZ	4	--

To block for order effects, the three control paradigm conditions were counterbalanced across the six subjects using two 3 x 3 Latin squares to create a unique order for each subject. To block for learning effects, the state of the LZ was randomized across the 16 unique LZs used. The instances of the automation error were also randomized.

Dependent Variables

The metrics of the evaluation were workload, control precision, and decision making. Additionally, qualitative feedback was gathered after each of the three control paradigm conditions and at the end of the experiment. The metrics are summarized in Table 7.

Table 7. Experimental metrics.

Dependent Variable	Description	Type	Frequency
Workload	Participant's self-ratings on NASA TLX workload scales	Numeric	Collected after each trial
Control Precision	Distance from the center of the actual landing location to the center of the LZ	Numeric	Only data from trials with correct landings
Efficiency	Time from start of trial to landing	Numeric	Only data from trials with correct landings
Decision Making	Correct land or abort decision	Categorical	Data collected for each trial
Decision Making	Distance from the LZ center when the final decision was made	Numeric	Data collected for each trial
Safety	Subjective ratings of safety on a Likert scale	Numeric	Data collected for each trial

The hypothesis can be defined in terms of the dependent variables, as shown in Table 8.

Table 8. The hypothesis of the protocol test.

Control Paradigm	Workload	Precision	Decision Making	Efficiency
(H) Human	High	Low	Higher correct decision rate Earlier decision time	Low
(A) Automation	Low	High	Lower correct decision rate Later decision time	High
(H + A) Human + Automation	Low	High	Higher correct decision rate Earlier decision time	High

Scenario

Participants conducted the final approach and landing portion of a flight. The scenario started with the vehicle in the air, just before the final waypoint, and then proceeded on to the final approach to the LZ. The scenario ended with either the vehicle landing or the vehicle aborting the landing. All landing scenarios were designed with similar trajectories and overall difficulty level. All trials started at an average altitude of 150 m ($SD = 0.3$ m) above ground and an average of 1038 m ($SD = 6.9$ m) from the LZ. The trajectory required a slight turn (average 6.9 degrees, $SD = 4.7$ degrees) to align with the final approach vector (to prevent the participants from seeing the LZ right away), and then a straight flight to the LZ. Nominal groundspeed to fly the trajectory was 50 km/h ($SD = 9.2$ km/h). The horizontal (birds-eye view) and vertical (altitude) profiles of the nominal trajectory are shown in Fig. 7.

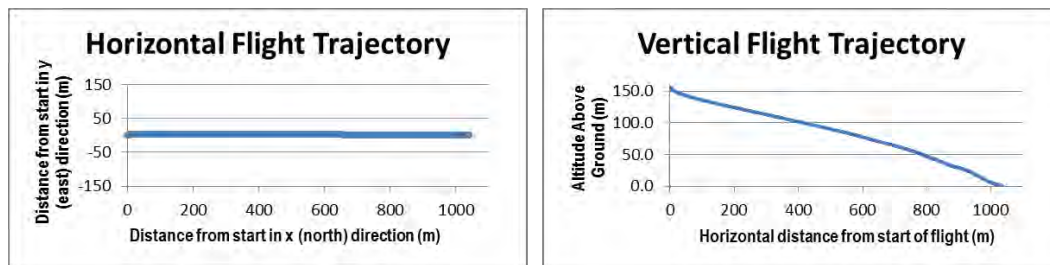


Figure 7. Horizontal and Vertical (altitude) flight plan trajectory of all trials.

Under the three control paradigms: 1) in the *Human (H)* condition, the participant informed an experimenter of the land/abort decision and controlled the vehicle with the steering wheel and pedals if the decision was to land; 2) in the *Automation (A)* condition, the vehicle made the final landing decision and the participant could not take control; and 3) in the *Human + Automation (H + A)* condition, the participant made the land/abort decision via the navigation display, and the automation controlled vehicle movement.

Simulation Flight Dynamics

The dynamic model of the vehicle was based on the controlled dynamics of an autonomous helicopter (the Boeing Unmanned Little Bird). This vehicle was chosen because it is a full-scale vehicle that has sufficient control loops that enables anybody to fly it. Additionally, this model is able to accept trajectories sent to the vehicle in an autonomous operation mode. The mathematical details of the model can be found in Scherer (2011). The sensors were simulated using the same real-time algorithms developed for an actual helicopter perception system tested in autonomous landing experiments (Scherer, Chamberlain, & Singh, 2012). The LZ assessment system takes as input a point cloud generated by sensor returns and evaluates this point cloud to determine if an area is landable by fitting a plane to a section of the space and determining the roughness of the terrain and the slope of the patch. If the slope or roughness is too large, an area is declared as not safe for landing (Scherer et al., 2012).

Procedure Section

After filling out initial paperwork and a demographics questionnaire, participants were briefed on the RAV concept, the HMI, the RAV Simulator, and the tasks that they would be performing. Participants were allowed to practice final approaches, landings, and aborts in the three control paradigm conditions, until they were comfortable in each. Participants were instructed that the automation was imperfect and could give an incorrect assessment of the LZ. Participants filled out a TLX survey after every trial. After all 16 trials in one control condition were completed, participants filled out a questionnaire. Finally, after all three conditions were tested, a final questionnaire was conducted.

Evaluation Results

Analyses of variance (ANOVA) using the F-test were conducted to explore the effects of the independent variables (control paradigm and LZ status) on the dependent variables. In cases where the F-test indicated a significant effect of an independent variable, the Least Significant Difference (LSD) post-hoc test was used to investigate which pairs of means differed. There were only two control paradigm conditions (H and H + A) where errors of omission or commission occurred. For the ANOVA, we averaged all measures for each condition for each participant to obtain a single value for each participant (one for each condition). Since the F-test with two groups is equivalent to the t-test, it was still used in these cases. Results are reported as significant for $\alpha = 0.05$ and marginally significant for $\alpha = 0.10$. Even though the number of participants was low (6), the effect size for all analysis was found to exceed Cohen's (1988) convention for a large effect ($d > .80$).

Workload

The NASA TLX workload tool (Hart & Staveland, 1988) was used as the metric for workload evaluation. This data was collected after each trial. A total of 286 data points were collected out of the expected 288, based on the experimental design in Table 7. With six workload subscales, each on a range of 1-10, the total workload range was 0–60. The total workload results (mean and standard error) are shown in the left side of Fig. 8 (an asterisk indicates a statistically significant

difference from other conditions). There was a significant difference in the average total workload between the three groups ($F(2,10) = 10.439, p = .003, d = 1.3$). This suggests a significant relationship between total workload and control paradigm. The average total workload for A was 8.6 ($SD = 5.0$), for H was 21.4 ($SD = 10.1$), and for H + A was 6.3 ($SD = 2.5$). A post-hoc LSD test shows that the average total workload for H was significantly higher than both A and H + A ($p = .005, d = 1.6$ and $p = 0.002, d = 2.1$, respectively). There was no significant difference between A and H + A. The results suggest that the H + A paradigm retains the workload benefits of the A paradigm.

Fig. 8 (right) shows the workload ratings (mean and standard error) along the six TLX subscales. Overall, the ratings show low to medium workload ratings on all scales since all ratings were below 5 on the 10-point scale. The H control paradigm had the highest rating for all workload variables except frustration, whereas the A paradigm had the highest rating. The average workload ratings for all subscales (mental, physical, temporal, performance, effort, and frustration) were significantly related to the control paradigm ($p < 0.05, d > 0.8$ in all cases). Results of LSD post-hoc tests detail the relationships between different control paradigms for the individual subscales.

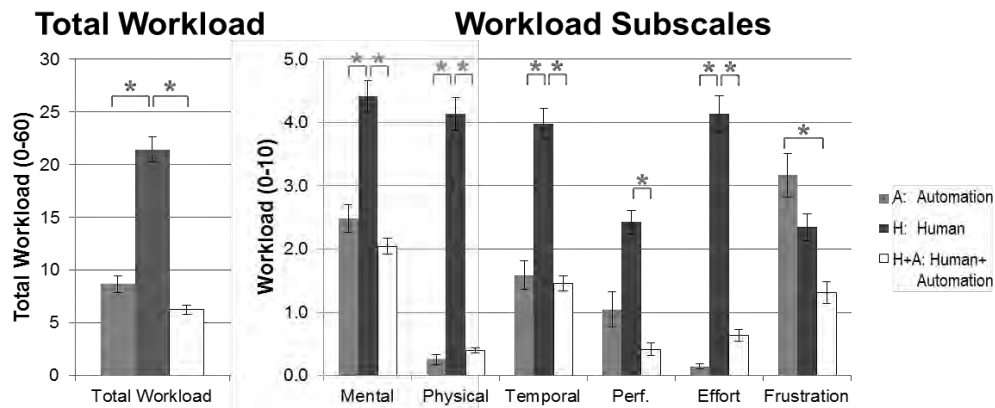


Figure 8. Total Workload ratings (left) and NASA TLX Workload Subscale ratings (right). (Perf. = Performance)

The average workload ratings for H were significantly different from H + A ($p < 0.05, d > 1.5$) in all the cases except frustration. The average workload ratings for A were significantly different from H ($p < 0.05, d > 1.0$) in all the cases except performance and frustration. The difference for performance was marginally significant ($p = .069, d = 1.0$). There was no significant difference between A and H + A for all subscales except frustration ($p = 0.01, d = 1.2$).

The results suggest that the H paradigm imposes the highest workload demand and A + H imposes the lowest for the majority of workload subscales. The high frustration rating in condition A could be due to the fact that the operator was unable to override the automation decisions in unsafe situations. Similarly, mental effort was required for all control paradigms by the very nature of the flying task. Even in the A and H + A paradigms, participants still monitored the automation and maintained situational awareness.

Precision

The metric used to assess landing precision was the distance from the center of the LZ to the actual landing spot. A total of 140 data points were collected out of the expected 144 based on the

experimental design in Table 7 (clear LZs). Fig. 9 (left) depicts the precision mean (and standard error). There was a significant effect of control paradigm on landing precision ($F(2,10) = 79.50$, $p < .001$, $d = 1.2$). The average precision score for A was 1.6 m ($SD = 0.2$), for H was 11.7 m ($SD = 2.8$), and for H + A was 1.4 m ($SD = 0.4$). The LSD post-hoc test indicates that H was significantly less precise than both A and H + A ($p < .001$). There was no statistically significant difference in precision between A and H + A. The results suggest that the Human + Automation paradigm retains the precision benefits of the Autonomous paradigm.

Efficiency

The time to fly the scenario from start to finish (landing only) was used as a metric for efficiency. A total of 140 data points were collected out of the expected 144, based on the experimental design shown in Table 7 (clear LZs). Fig. 9 (right) depicts the mean (and standard error) efficiency. The average efficiency score for A was 66.1 s ($SD = 0.1$), for H was 94.2 s ($SD = 25.0$), and for H + A was 64.7 s ($SD = 2.9$). The control paradigm significantly affected efficiency ($F(2,10) = 7.09$, $p = 0.012$, $d = 1.3$). H was significantly higher than both A and H + A ($p = 0.01$ and $p = 0.008$ respectively) based on the LSD post-hoc test. There was no significant difference in task time between A and A + H. The results suggest that the Human + Automation paradigm retains the efficiency benefits of the Autonomous paradigm.

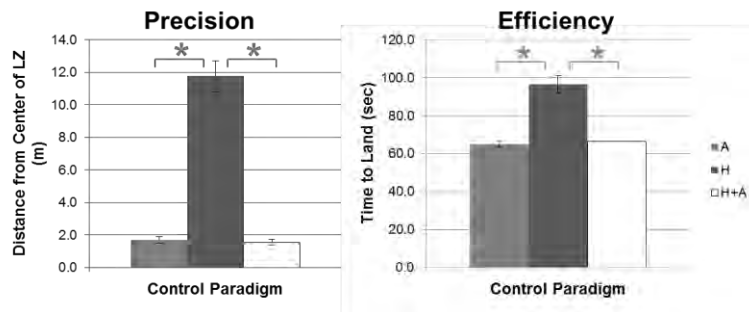


Figure 9. Landing precision and efficiency.

Decision Making

Decision-making data was collected for each trial. A total of 286 data points were collected out of the expected 288, based on the experimental design in Table 7. Two metrics were used to assess decision making in the evaluation scenario. The first metric was the error rate in difficult situations. This was calculated as the number of incorrect land or abort decisions divided by the number of trials for each experimental condition. It is important to note that the automation errors were predetermined “errors” specifically triggered for experimental design purposes, and thus would not be representative of the error rate in a fielded system under all operation types. This was done in order to explore challenging situations. There were two types of automation errors:

- Scanner unable to detect unstable landing surfaces;
- Automation errors of commission and omission (unable to detect an obstacle or incorrectly classifying a safe LZ as unsafe).

The error analysis was based on the categorized frequency of errors table and not a comparison of direct measures for each trial between the groups, and so the Fisher Exact Test was used to analyze the error results. The error rate results (mean and standard error) are shown in Fig. 10 (left). The error rate for A (37.9%) was significantly ($p < 0.001$) higher than both H (1.0%) and

H + A (3.2%). Thus, human operators were able to compensate for those situations where the

automation was not designed to prevent errors or where the automation failed to correctly scan the LZ. There was no significant difference in error rate between H and H + A. The results suggest that the Human + Automation paradigm retains the decision-making benefits of the Human-only paradigm.

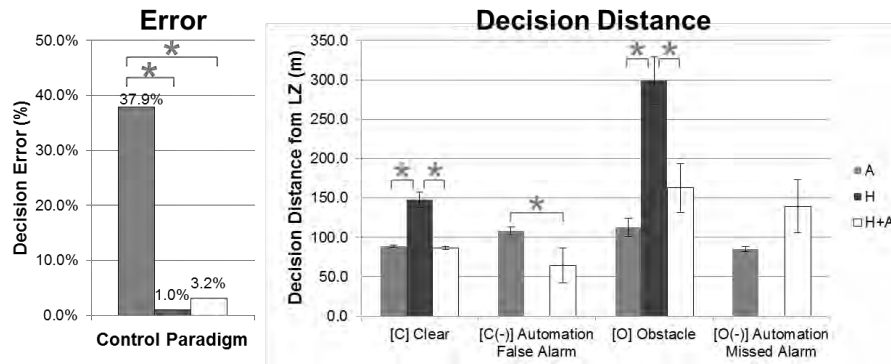


Figure 10. Decision making: Errors and distance from LZ when decisions were made.

The second metric used to assess decision making was the distance from the LZ center when the land/abort decision was made. A total of 284 data points were collected out of 288 expected, based on Table 7. The results are shown in Fig. 10 (right). In the H paradigm, humans aborted earlier ($M = 299$ m, $SD = 96.7$) when an obstacle was present than when the LZ was clear ($M = 150$ m, $SD = 27.1$). The difference is marginally significant ($F(1,5) = 4.13$, $p = .098$, $d = 2.1$). In essence, they aborted the moment that they were sure of an obstacle.

For a clear LZ ([C] in Fig. 10), the control paradigm had a significant effect on participants' decision making ($F(2,10) = 24.05$, $p < .001$, $d = 1.4$). A LSD post-hoc analysis indicated that participants waited significantly longer ($p < .001$, $d = 3.2$) in the A ($M = 89$ m, $SD = 3.4$) and H + A ($M = 65.2$ m, $SD = 22.8$, $p < 0.001$, $d = 3.4$) than in the H paradigm ($M = 150$ m, $SD = 27.1$), in order to take advantage of the scanning system to corroborate their own assessment before making the final decision. The difference between H + A and A was not statistically significant.

The decision distance for A ($M = 112.5$ m, $SD = 27.5$) was limited by the sensor range. When automation was incorrect for a clear LZ ([C(-)] in Fig. 10, participants in H + A paradigm ($M = 60.5$ m, $SD = 38.2$) took time to override the automation in every case. This differed significantly ($F(1,5) = 2.5934$, $p = 0.044$, $d = 1.7$) from the A paradigm ($M = 108.3$ m, $SD = 12.4$), where the automation aborted automatically.

The control paradigm again had a significant effect ($F(2,10) = 6.8039$, $p = .014$, $d = 1.4$) on participants' decision making when the LZ was obstructed ([O] in Fig. 10). Similarly, a post-hoc analysis indicated that participants waited significantly longer ($p = 0.021$, $d = 1.2$) in the A ($M = 112$ m, $SD = 27.5$) and H + A ($M = 162.7$ m, $SD = 123.3$) than in the H paradigm ($M = 299$ m, $SD = 96.7$). The difference between H + A and A was not statistically significant. In the H + A condition, the participants all detected the obstacle but took different strategies with relation to the automation. Three of the six participants aborted immediately (not waiting for the automation), and three delayed their decision in order to take advantage of the scanning system to corroborate their own assessment before making the final decision.

When the automation failed to detect an obstructed LZ ([O(-)] in Fig. 10, the participants in H + A ($M = 139.5$, $SD = 83.4$) behaved much in the same way for the correct detection of an obstruction case ("[O]"). The participants all detected the obstacle but took different strategies with relation to the automation. Three of the six participants aborted immediately, not waiting for

the automation. The remaining three participants waited for confirmation from the automation.

When it failed to sound the alarm, they then aborted the landing. The average decision distance in H + A did not differ significantly from the A paradigm ($M = 84.9$ m, $SD = 7.6$), when humans were helpless to override the automation, even if they knew it was in error.

Overall, the results suggest that Human + Automation retains the decision-making benefits of the Human paradigm when are obstacles are present and uses automation to confirm a clear LZ.

Safety Rating

Subjects were asked at the end of each of the 48 trials completed if they felt safe during the trial on a scale of 1 (unsafe) to 5 (safe). A total of 285 data points were collected out of the expected 288, based on the experimental design in Table 7. The results of subject ratings are shown in Fig. 11. The average safety rating for A was 3.5 ($SD = 0.6$), for H was 4.2 ($SD = 0.5$), and for H + A was 4.5 ($SD = 0.6$). The control paradigm significantly affected safety ($F(2,10) = 13.0808$, $p = .002$, $d = 1.4$). The A paradigm was significantly rated as less safe than both H and H + A based on the LSD post-hoc test ($p = .005$, $d = 1.3$, and $p < .001$, $d = 1.7$). Additionally, there was no significant difference in safety rating rate between H and A + H. The results suggest that participants felt safer when they had some form of control.

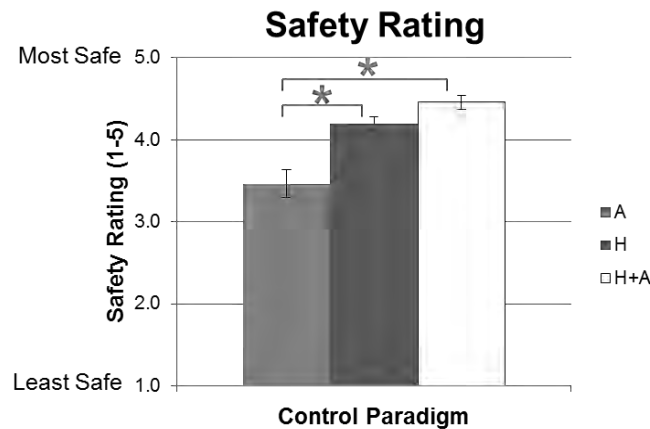


Figure 11. Subjective safety ratings.

Discussion

Overall, the results confirmed the experiment hypotheses (Table 9). For all of the evaluation metrics, the results suggested that the Human + Automation paradigm leveraged the benefits of either Human-only or Automation and was not significantly affected by any drawbacks associated with either paradigm.

Subject comments showed strong support for the H + A paradigm:

- “Automation is very good; Automation + Human is even better.”
- “Shared autonomy was the clear winner.”
- “I like the hybrid [paradigm]—it is more efficient and safe.”
- “The Human + Autonomous was fast to land but still allowed me to step in if the vehicle was going to do something unsafe.”

Table 9. Confirmed Experiment Hypothesis, where ✓ indicates best condition per metric.

Control Paradigm	Workload	Precision	Decision Making	Efficiency
Human	Poor	Poor	Good	Poor
Automation	Good	Good	Poor	Good
Human + Automation	✓ Good	✓ Good	✓ Good	✓ Good

The results also validated the motivation for a joint H + A design concept for the RAV. Some H + A paradigm strengths, as identified by the subjects, are as follows:

- Easy decision making—choice between land or abort;
- LZ classification through color changes: magenta (LZ found and being assessed), green (LZ validated as safe), and red (LZ assessed as unsafe);
- Allows pilot input: the pilot confirms the safety of landing decisions;
- Allows the pilot to focus on tasks other than flying;
- The system behavior, including default behavior, is acceptable.

While not measured directly, the decision-making results imply that the situation awareness of the operator played a crucial role in the landing or aborting decision. When both the operators and the automation were given permission to land or abort the takeoff, the operators leveraged their situational awareness of the environment to abort landings that they deemed were unsafe. More importantly, operators were able to use their awareness of the environment to compensate for instances where the automation was in error.

Conclusion

The goal of the RAV HMI presented here is to support highly safe and reliable operations by supporting human involvement in the control of the RAV at whatever level of detail is necessary for optimal achievement of mission objectives. The HMI design utilizes multiple modes of control, allowing the operator a span of control from tactical to strategic, with varying attention demands.

The role of human operators interacting with highly capable, complex automation continues to be an area of fertile research. This paradigm pushes the requirements of human-robot mixed-initiative control by enabling a human with no pilot training to control a highly automated RAV in the way humans often want to be integrated—*flexibly in charge and aware but not required to issue every command*. It is currently impractical to think that a fully autonomous vehicle can be built that has the safety margins necessary to carry humans onboard, while still being affordable enough to see widespread deployment. Mixed-initiative control embraces the fact that there is a human operator onboard; accordingly, we designed the HMI to leverage the strengths of an in situ operator while mitigating the lack of pilot training. The in situ operator has tremendous situation awareness that can be brought to bear to make the crucial decisions. As the evaluation showed, a mixed-initiative control solution can leverage the decision-making strengths of human operators with the precision, efficiency, and workload strengths of automation.

As the perception and planning systems mature, future HMI research will focus on how to integrate the wealth of information and functionality available to enhance the operator's situation awareness without overloading him or her with information. The perception system, for instance, can “paint” the terrain with measures of “goodness”—how much of this information should be displayed to the operator? Likewise, the planning system can generate a family of safe approach

paths, landing sites, and abort paths—should the operator have a role in deciding which paths to follow?

In future work, as the perception systems, the automation, and the HMI mature, increased functionality has implications for the joint human-automation control. For instance, human control can be expanded by enabling a real-time change in the prior designated LZs. Enhanced automation control may be enabled with new, longer-range sensors. Yet as capabilities grow, the correct way to integrate the shared control is not clear. Given that the evaluation was conducted in a virtual environment, more research is needed to assess the level of acceptance of automated tasks by an operator in a real RAV. The operator's capabilities and limitations can be further characterized to understand both the risk and the benefits a human operator can bring, along with subsequent implications for design, operations, and training.

Acknowledgements

This research was supported by a contract with the Defense Advanced Research Projects Agency (DARPA) and the U.S. Air Force under award no. FA8650-10-C-7073. The opinions expressed herein are those of the authors and do not necessarily reflect the views of DARPA or the U.S. Air Force. Additionally, the authors would like to acknowledge the efforts of the research and development team that supported the development of the prototypes, the running of the experiments, and the analysis of the data, which includes Trent Reusser, Robert De Mers, Matthew Aasted, Sankalp Arora, and Hugh Cover. Finally, the authors wish to thank Leslie Potter for review of the manuscript.

References

- Abbott, T. (1993). *Functional categories for future flight deck designs*. NASA Technical Memorandum TM-109005. Hampton, VA: NASA Langley Research Center.
- Allen, J. F. (1999). Mixed-initiative interaction. *IEEE Intelligent Systems*, 14(5), 14-16.
- Alter, K. W., Erickson, J. B., Goins, R. T., Hofer, E. F., Koehn, W. L., Miles, W. L., . . . Pfaff, T. A. (1995). *High speed research flight deck design and integration flight deck concepts*. Seattle, WA: Boeing/McDonnell Douglas Industry Team.
- Barnard Microsystems Limited (2011). Reliability of unmanned aircraft. Retrieved from <http://www.barnardmicrosystems.com>
- Beringer, D. B. (2002). Applying performance controlled systems, fuzzy logic, and fly-by-wire controls to general aviation. Tech. Rep. DOT/FAA/AM-02/7. Oklahoma City, OK: Federal Aviation Administration.
- Billings, C. E. (1997). *Aviation automation: The search for a human centered approach*. Mahwah, NJ: Erlbaum.
- Carbonell, J. R. (1970). Mixed-initiative man-computer instructional dialogues. Final report. Tech. Rep. No. BBN-1971, Job No. 11399. Cambridge, MA: Bolt Beranek & Newman, Inc.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- DARPA (2010). Transformer (TX) vertical takeoff and landing roadable air vehicle. Retrieved from: https://www.fbo.gov/download/d10/d10015bd49ec44f4414d65635bc1a37e/TX_BAA_Version_62_&_Appendix_A.pdf
- Donaldson, T. & Cohen, R. (1997). A constraint satisfaction framework for managing mixed-initiative discourse. In *Proceedings of the AAAI Spring Symposium: Computational Models for Mixed Initiative Interaction*, Palo Alto, CA: AAAI Tech. Rep. SS-97-04, pp. 37-43.

- FAA (1988). System design and analysis. Federal Aviation Administration Advisory Circular AC 25.1309-1A. Retrieved from http://www.faa.gov/documentLibrary/media/Advisory_Circular/AC%2025.1309-1A.pdf
- Flemisch, F., Adams, C. A., Conway, S. R., Goodrich, K. H., Palmer, M. T., & Schutte, P. C. (2003). The H-Metaphor as a guideline for vehicle automation and interaction, Tech. Rep. No. NASA/TM—2003-212672. Hampton, VA: NASA, Langley Research Center.
- Flemisch, F., Heesen, M., Hesse, T., Kelsch, J., Schieben, A., & Beller, J. (2012). Towards a dynamic balance between humans and automation: Authority, ability, responsibility and control in shared and cooperative control situations. *Cognition, Technology & Work*, 14(1), 3-18.
- Genta, G., Morello, L., Cavallino, F., & Filtri, L. (2014). *The motor car*. Springer: The Netherlands.
- Goodrich, M. A., Olsen, D. R., Crandall, J. W., & Palmer, T. J. (2001). Experiments in adjustable autonomy. In *Proceedings of IJCAI Workshop on Autonomy, Delegation and Control: Interacting with Intelligent Agents*. Tucson, AZ, pp. 1624-1629.
- Hart, S. G., & Staveland, L. E. (1988). Development of a multi-dimensional workload rating scale: Results of empirical and theoretical research. In P. Hancock & N. Meshkati (Eds.), *Human mental workload* (pp. 139-183). North Holland, The Netherlands: Elsevier.
- Jump, M., Perfect, G. D., White, M. D., Floreano, D., Fua, P., Zufferey, J. C., . . . Bühlhoff, H. H. (2011). myCopter: Enabling technologies for personal air transport systems. *The future rotorcraft—enabling capability through the application of technology*. London, UK.
- Kortenkamp, D., Bonasso, R. P., Ryan, D., & Schreckenghost, D. (1997). Traded control with autonomous robots as mixed initiative interaction. In *Proceedings of the AAAI Symposium on Mixed Initiative Interaction*. Stanford, CA.
- Landén, D., Heintz, F., & Doherty, P. (2012). Complex task allocation in mixed-initiative delegation: A UAV case study. In N. Desai, A. Liu, & M. Winikoff (Eds.), *Principles and practice of multi-agent systems* (pp. 288-303). Berlin: Springer.
- Lee, S. M., Kim, S. Y., & Feigh, K. M. (2009). Structural framework for performance-based assessment of ATM systems. In *Proceedings of the AIAA Conference on Aviation Technology, Information, and Operations*. Hilton Head, SC.
- Miller, C. A., Funk, H. B., Dorneich, M., & Whitlow, S. D. (2002, October). A playbook interface for mixed initiative control of multiple unmanned vehicle teams. In *Proceedings of the 21st Annual Meeting of the Digital Avionics Systems Conference*. Irvine, CA.
- Miller, C. A., & Parasuraman, R. (2007). Designing for flexible interaction between humans and automation: Delegation interfaces for supervisory control. *Human Factors*, 49(1), 57-75.
- Moore, M. D. (2006). The third wave of aeronautics: On-demand mobility. In *SAE General Aviation Technology Conference and Exhibition*, SAE (pp. 01-2429).
- Pacaux-Lemoine, M-P., Debernard, S. (2000). A common work space to support the air traffic control. *Control Engineering Practice*, 10 (pp. 571-576).
- PAL-V (2012). Flying car makes successful maiden flight. Retrieved from http://pal-v.com/wp-content/uploads/2012/03/Pal-V_press_release.pdf
- Parasuraman, R., & Bowers, J.C. (1987). Attention and vigilance in human-computer interaction. In A. Gale & B. Christie (Eds.), *Psychophysiology of the electronic workplace* (pp. 163-194). London: Wiley.

- Parasuraman, R. and Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. *Human Factors*, 39(2), 230-253.
- Parasuraman, R., Sheridan, T.B., & Wickens, C.D. (2000). A model for types and levels of human interaction with automation. *IEEE Transaction on Systems, Man, and Cybernetics—Part A: Systems and Humans*, 30(3).
- Rasmussen, J., & Vicente, K. J. (1989). Coping with human errors through system design: Implications for ecological interface design. *International Journal of Man-Machine Studies*, 31(5), 517-534.
- Reason, J. (1990). *Human error*. New York: Cambridge University Press.
- Riley, V. (1989). A general model of mixed-initiative human-machine systems. In *Proceedings of the Human Factors and Ergonomics Society (HFES) Conference*. Denver, CO.
- Rognin L., Salembier P., & Zouinar M. (2000). Cooperation, reliability of socio-technical systems and allocation of function. *International Journal of Human-Computer Studies*, 52, 357-379.
- Sheridan, T. (1987). Supervisory control. In G. Salvendy (Ed.), *Handbook of Human Factors* (pp. 1244-1268). New York: John Wiley & Sons.
- Scherer, S. (2011, May). *Low-altitude operation of unmanned rotorcraft*. (Doctoral dissertation). Tech. Rep. CMU-RI-TR-11-03, Robotics Institute, Carnegie Mellon University. Retrieved from https://www.ri.cmu.edu/pub_files/2011/5/thesis-sebastian-final.pdf
- Scherer, S., Chamberlain, L., & Singh, S. (2012). Autonomous landing at unprepared sites by a full-scale helicopter. *Robotics and Autonomous Systems*, 60(12), 1545-1562.
- Schmidt, K. (1990). *Analysis of cooperative work: A conceptual framework*. Risø National Laboratory, DK-4000 Roskilde, Denmark (Risø-M-2890).
- Smith, P. J., McCoy, E. C., Layton, C. (1997). Brittleness in the design of cooperative problem-solving systems: The effects of user performance. *IEEE Transactions on Systems, Man, and Cybernetics—Part A*, 27(3).
- Terrafugia (2012, July). Terrafugia's Transition® street legal airplane continues flight and drive testing. Retrieved from http://www.terrafrugia.com/news_media.html.
- Weissler, P., Mason, D., & Woock, K. (2012). Flying car makes debut in New York. *Benefits*, 2011, 10-17.
- Woods, D. (1996). Decomposing automation: Apparent simplicity, real complexity. In R. Parasuraman & M. Mouloua (Eds.), *Automation and human performance: Theory and applications* (pp. 3-17). Hillsdale, NJ: Lawrence Erlbaum Associates.

Authors' names and contact information: M. C. Dorneich, Department of Industrial and Manufacturing Systems Engineering, Iowa State University, Ames, IA, USA. Email: dorneich@iastate.edu; E. Letsu-Dake, Honeywell Laboratories, Golden Valley, MN USA. Email: Emmanuel.Letsu-Dake@honeywell.com; S. Singh, Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, USA. Email: ssingh@cmu.edu; S. Scherer, Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, USA. Email: basti@andrew.cmu.edu; L. Chamberlain, Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, USA. Email: lylesjazz@gmail.com; M. Bergerman, Robotics Institute, Carnegie Mellon University, Pittsburgh, PA, USA. Email: marcelbergerman@gmail.com.