International Journal of Social Robotics (2021) 13:1761–1776
https://doi.org/10.1007/s12369-021-00757-x

Utility of Functional Transparency and Usability in UAV Supervisory Control Interface Design

Wenjuan Zhang[1] · **David Feltner**[2] · **David Kaber**[3] · **James Shirley**[4]

Accepted: 28 January 2021 / Published online: 23 February 2021

Abstract

A basic notion of transparency in automated systems design is the need to support user tracking and understanding of system states. Many usability principles for complex systems design implicitly target the concept of transparency. In this study, we made comparison of a "baseline" control interface mimicking an existing available UAV ground control station with an "enhanced" interface designed with improved functional transparency and usability, and a "degraded" interface which removed important design features. Each participant was extensively trained in the use of one of the interfaces and all simulated UAV control tasks. Each participant was tested in four trials of a typical military concept of UAV operation with different mission maps and vehicle speeds. Results revealed participants using the enhanced interface to produce significantly faster task completion times and greater accuracy across all UAV control tasks. The enhanced features were also found to promote operator understanding of the system and mitigate workload. By defining and setting automation transparency as an overarching design objective and identifying specific transparency and usability issues within existing GCS designs, we weress able to design and prototype an enhanced interface that more effectively supported human-automation interaction. Automation transparency as a high-level design objective may be useful for expert designers; whereas, usability design guidelines, as "building blocks" to transparency, may be a useful tool for new system designers.

Keywords Functional transparency · Usability · Unmanned aerial vehicles · Supervisory control · Interface design

1 Introduction

In a report from the Air Force Research Laboratory in 2009, it was observed that Unmanned aerial vehicle (UAV) accidents had a substantially higher rate of occurrence as compared to general Air Force aviation mishaps [1]. Although riskier missions may have contributed to the higher incident rate for UAVs, the literature suggests that approximately 60% of UAV accidents have been attributed to human factors issues in vehicle control [2, 3] and 30% of accidents have been specifically associated with designs of interface technology [4]. This situation is concerning because UAV applications are expanding from military uses (e.g., information collection, precision strikes) to more civilian uses (e.g., law enforcement, first response, package delivery). Furthermore, billions of dollars have already been invested in research efforts to improve UAVs in terms of hardware, software, and human-system interactions [5].

In general, as "unmanned" systems become more autonomous, control interfaces are used more for monitoring and diagnosing of system faults, and less for direct manipulation [6]. The trend of increasing level of automation (LOA) has also raised several issues in unmanned systems interface design. UAV operators have complained of a lack of feedback on control inputs and difficulty in detecting and correcting errors [7, 8]. This situation is problematic because pilots and payload controllers require a substantial amount of information from vehicles, sensors and other observers in order to ensure the health and welfare of a UAV and to achieve mission goals. However, human operators have a

✉ Wenjuan Zhang
wzhang28@ncsu.edu

✉ David Kaber
dkaber@ise.ufl.edu

1 Dataminr, New York, NY, USA

2 United States Military Academy, West Point, NY, USA

3 Department of Industrial & Systems Engineering, University of Florida, P.O. Box 116595, Gainesville, FL 32611-6595, USA

4 NAVAIR, New Bern, NC, USA

finite capacity for information processing. This limitation is manifested when the operators are under workload, which is impacted by display content and configuration. Other UAV operator complaints documented in the literature include lack of display design consistency and non-intuitive symbology and visually fatiguing use of color [2, 9, 10]. Given the above issues, ensuring that the "right" information is available for operators for specific tasks, and that it is presented in a clear and consistent manner, is critical to mission success.

1.1 The Concept of Automation Transparency

Historical human factors research has [11, 11] identified several challenges in the context of human interaction with automated agents, including understanding the current system state, comprehending its reasons for behavior, and projecting the next system behavior. Similarly, Selkowitz et al. [12] contended that it is critical for human operators to have an accurate understanding of an agent's actions, operating environment, reasoning, projections and states of uncertainty to effectively exploit system functionality. This contention has been identified as the concept of automation transparency and is garnering more and more attention as some unmanned systems approach autonomous operation in certain environments.

Related to Selkowitz et al. [12] work, Chen et al. [13] proposed a model of agent transparency to support operator situation awareness (SA) of a mission environment through the agent. In Chen's model, which is known as the SA-based Agent Transparency (SAT) model, transparency is defined as a descriptive quality of an interface, based on its capabilities to afford a user comprehension of an intelligent agent's intent, performance, future plans, and reasoning process in context. Chen et al. [13] suggested that automated agent interfaces should present three levels of information. Level 1 transparency provides basic details about the agent or the task, such as current system status and route information. Level 2 information entails reasoning and can be conveyed through a representation of resource limitations, constraints/affordances, feasibility, risk, and trade-offs between alternatives to help an operator make good decisions for a mission. Level 3 information provides the operator with an understanding of the expected outcomes, such as time-to-complete objectives and the probability of success for certain goals.

1.2 Design Guidelines for Transparency

With a focus on the need for transparency, a number of studies have formulated general design guidelines for "automation presentation". Lyons [14] suggested that automation should provide rationales for courses of action so a human knows why a system is doing what it is doing. Others have suggested that operators should be provided with guidance on different modes of system operation and decision making in real-time [15]. Still other concepts of automation transparency are very specific and focus on a need to support user trust in a system and to realize appropriate trust calibration through various types of display cues [16]. Associated recommendations have included identifying and highlighting automation errors during operation [17] or presenting uncertainty, such as confidence levels in specific system states and predicted outcomes Chen et al. [13].

Kilgore and Voshell [18] made explicit suggestions on improving transparency in unmanned systems within the framework of ecological design. One recommendation was to increase the perceptual availability of task-critical information by presenting key system variables that affect automated outcomes, or coding map features using visual cues such as shape, size, and opacity. Another recommendation was to present information in context. For example, health-and-status values should be provided against the frame of expected values or nominal ranges. The third recommendation was to manage operator attention. Display figures should automatically direct an operator's attention to critical system process information and cause less critical information to recede into the background.

Selkowitz et al. [12] also summarized display design techniques that are best suited for presenting autonomous system state information to support transparency in communication interfaces. In general, these techniques included ecological interface design, integrated displays, and pre-attentive cuing. Some specific examples included the use of metaphor-based presentations of information, integration of multiple pieces of information into a single icon (like glyphs), and use of shading, color, or size coding to promote effortless and quick processing of information. The authors also conducted a few experiments to test the effectiveness of these techniques for supporting transparency [19, 20]. Overall, the identified features were found to be helpful for improving transparency, aiding operators in system monitoring, maintaining a high level of operator situation awareness, gaining operator trust, and improving operator performance without a cost of increased workload. Similar findings have also been reported by other studies [21–23].

Although not made explicit, most of the above recommendations for automation transparency use two mechanisms, including implementing functionalities that are designed to promote transparency and improving usability of an interface. Here we provide a few examples, to demonstrate that these two mechanism are not independent in practice. The first example comes from Study 1 of Selkowitz et al. [12]. The autonomous robot display used green triangles to indicate beneficial terrains and red circles to mark ambush hazards. This interface required the robot to be equipped with a function for detecting surrounding terrain and analyzing its beneficial vs. hazardous nature. This functionality was designed to promote transparency and was also accompanied with usability by appropriately color-coding terrain information to help the operator process complex information. The second example comes from the ecological fuel management display recommended by Kilgore and Voshell [18]. Their fuel management display depicted fuel consumption for a mission in the context of overall fuel capacity. The overall capacity measure was dependent on the amount of currently available fuel, anticipated fuel reserves, and the minimum amount of remaining reserve fuel to ensure mission safety. This fuel management tool required extensive development of complex computing and analysis functionalities. In addition, the fuel management display was designed with usability, including appropriate use of color and shapes to represent the various aspects of the overall system capacity.

In summary, implementing additional functions (e.g., pre-attentive cueing on dynamic parameters and/or potential system outcomes) can improve automation transparency by aiding operators with system comprehension and future state projection. On the other hand, improving interface usability allows operators to effectively perceive information with less time and effort.

1.3 Research Motivation

Based on a brief survey of the literature, a number of studies have investigated UAV interface design and implications. Some studies have looked at the number of UAVs that can be controlled by an operator at any one time [24–26]. Other studies have manipulated interface features in attempts to better support operator capabilities to manage multiple UAVs simultaneously [4]. However, many of these investigations have been limited in focus to the type of cameras, depth cues, and multi-modal displays integrated in system designs. Few studies have conducted in-depth examinations of automation transparency in UAV interface design. A literature review by Prewett [27] also indicated a lack of studies investigating what information components to present in UAV displays and how to organize information for optimal visual presentation.

The present research made use of automation interface design recommendations, presented across several of these studies, to further investigate the utility of functional transparency and usability features in UAV supervisory control interfaces. Given the relatively high and persistent accident rates with such systems, we consider it necessary for the human factors research community to validate principles for transparency in unmanned system interfaces and to support "enhanced" design of control interfaces for safe mission performance. In general, we expected that interfaces with functions designed for promoting transparency and enhanced usability would improve operator situation awareness, their performance on generic UAV tasks, and reduce operator workload. On the other hand, interfaces lacking such features were expected to be detrimental to user performance and workload responses.

2 Method

A simulator-based experiment was conducted to assess the above expectations implementing functional transparency and usability design guidelines in UAV control interface design. The study was designed to provide further evidence of the utility of specific design principles for supporting performance and moderating operator workload.

2.1 Participants

Forty-eight (48) participants (25 male, 23 females; age: 24.8 ± 3.9) were recruited from the NC State University campus population and surrounding community. All participants were required to: (1) have 20/20 or corrected vision and no color-impairment due to the study requirement for visual perception of colored objects/text on a display screen; (2) have general familiarity with the use of computers, as the study involved using an interface simulation presented on a desktop/laptop computer; and (3) have no experience with UAV supervisory control operations so as not to bias performance in the simulated control tasks. (All participants were provided with extensive training on a UAV control interface and in common vehicle control tasks. More information will be provided in Sect. 2.7.)

2.2 Tasks and Apparatus

During the experiment, participants completed a simulated UAV surveillance mission involving tasks representative of real-world fundamental control operations. Interface complexity and related hardware control varies based on UAV scale. Peschel and Murphy [28] classified UAV systems into four different categories, including micro, small, medium altitude long endurance (MALE), and high altitude long endurance (HALE). MALE and HALE UASs tend to be software-based (e.g., menus, point-and-click, touch, etc.), while the micro and small interfaces tend to be hardware-based (e.g., joystick, buttons, toggle switches, etc.) [28]. Despite reservations with a taxonomy of vehicles based on relative size classifications not being a crisp design reference, in order to investigate the effect of automation transparency in sophisticated automated systems, we created prototypes (using Just in Mind, San Francisco, USA) for a software-based UAV interface. The basic prototype interface design was referenced from a commercially available UAV ground control station (GCS) software, the Mission Planner (MP) by the ArduPilot Development Team. The Mission Planner interface was used as a reference because it includes most UAV interface components described in the current interface research [29–36]. Other commercial UAV interfaces typically use much simpler interface designs and focus on the camera view. A range of interface features were manipulated in prototypes for this study.

A surveillance mission scenario was designed for this experiment. This mission scenario was drafted by a military officer with extensive experience in authoring UAV mission orders with the objective of representing generic operation tasks. The tasks were also selected so that as many of the UAV interface components were used in support of the mission. Each mission consists of five tasks, including determining coordinates of specified objects of interest on a navigation display (i.e., coordinate task), determining the distance between two objects (i.e., distance estimation task), handling in-flight alerts and alarms (i.e., alarm fix task), prioritizing in-flight alerts and alarms (i.e., alarm prioritization task), and monitoring system status parameters and report anomalies (i.e., parameter warning detection task). The order of these tasks in each mission was arbitrary but the tasks were more or less evenly distributed across the timeline of each mission. The alarm messages presented in the alarm fix task, and the actions for resolution, are presented in Table 1. In each alarm fix task, only one alarm was presented and the participants were asked to choose the correct response action among all possible actions listed in Table 1. The alarm messages used in the alarm prioritization task and their priority levels are presented in Table 2. In each alarm prioritization task, three alarms, including one from each priority level, were presented. Participants were asked to report the correct priority level for each alarm. These alarm examples were selected to simulate common alarms in aviation while simplifying their priorities and solutions to avoid the need for extensive training as part of the experiment.

Table 1 Alarms used in Alarm Fix task

Alarms	Actions to fix alarm
Excessive rate of descent	Slow descent
Crash imminent	Pull up
Engine fire	Extinguish
Abort takeoff	Abort
Landing gear malfunction	Reset gear
Exited selected altitude	Change altitude
Bank angle > 35 degrees	Reduce angle
Autopilot disconnected	Reconnect
Suboptimal glide slope	Change slope
Maximum airspeed reached	Reduce speed
20% fuel remaining	Refuel

Table 2 Alarms used in Alarm Prioritization task

Priority level	Alarms
High Priority	Excessive rate of descent
	Stall impending
	Crash imminent
	Engine fire alarm
	Abort take off
	Landing gear failure
Medium Priority	Exit selected altitude
	Bank angle greater than 35 degrees
	Autopilot disconnected
	Suboptimal glide slope
	Maximum air speed reached
	20% fuel remaining
Low Priority	Quarterly service due
	Parking brake released
	Aircraft in flight
	Engine valve open
	Navigation active
	Instrument panel switch activated

Participants performed these tasks using design variations on the MP interface as presented on a standard computer monitor (Fig. 1). They were also provided with a mission description document that presented an overview of the vehicle trajectory as part of the mission. This document also outlined the "scheme of maneuver" for the vehicle, which listed operator control tasks for achieving the mission goals. This mission description remained viewable for participants throughout experiment trials and was frequently used for recalling tasks (also see Fig. 1).

In addition to the "scheme of maneuver" tasks, participants were posed with dynamic knowledge queries (DKQ) during the test trials that required them to answer questions based on the current status of the vehicle or mission scenario. The DKQs were designed to test participant knowledge of basic information about the UAV interface and

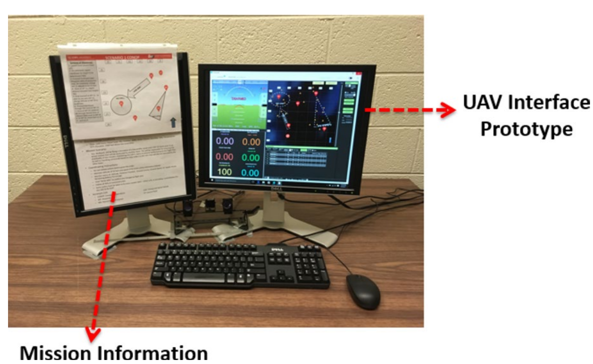

Fig. 1 Experiment apparatus (mission document on left; interface prototype on right)

Table 3 Example DKQs

Example DKQs
What is the name of the current flight axis?
What is the normal range for UAV altitude?
Under the Map Action function, how many "Delete" options are available?
What Northing is Waypoint 12 closest to?
What is your current vehicle ground speed?
What is your current completion percentage for this mission?
How many targets are inside [the named area of interest]?

mission, as well as their perception and comprehension of states of the vehicle, as potentially influenced by the interface design features and mission/task workload. Table 3 presents examples of DKQs relavent to UAV control tasks. The DKQs used in the experiment were slight variations on these examples (e.g., the current altitude instead of the current ground speed).

2.3 Interface Design Variations

Three different interface prototypes were developed and tested in this experiment. A baseline interface (Fig. 2) was created to most closely represent the functionality currently available in the original MP interface. This design was intended to represent a commercially available control interface. It is important to note that features without relevance to the experiment tasks were simplified or removed from the prototype. For example, the complex software and hardware configuration options in the original MP interface were not included in the baseline interface. Another example is that the original MP interface allowed the user to choose several different types of displays for viewing flight parameter data, such as digital number displays, analog displays, or data logging displays. To simplify and streamline the experiment design, the baseline interface only included the digital number displays, as the design represented the greatest degree of usability. Other features (e.g., alarms) were integrated to the interface to allow for investigation of common UAV control operations not currently supported by the MP.

For our research purpose, two other prototypes were created from the baseline interface: an enhanced interface and a degraded. Guidelines from the automation transparency literature [12, 18] were applied to create the interface variations. The enhanced interface followed many design guidelines to facilitate the predefined tasks, while the degraded interface violated specific guidelines without preventing operators from completing these tasks.

Figure 3 shows the enhanced interface with functionalities supporting transparency and usability improvements. The enhancements of the control interface focused on automation assistance and operator attention management. For example, the enhanced interface implemented automation for estimating distance between two objects (i.e., the distance estimation task). It also provided greater contextual information for individual menu options by grouping items into relevant categories. In addition, the enhanced interface applied pre-attentive cueing by color coding alarm priorities, highlighting suggested alarm solutions, introducing visual differentiation between past and future waypoints, and alerting abnormal flight parameters. The interface manipulations are summarized in Table 4. These features were designed to facilitate task operation and

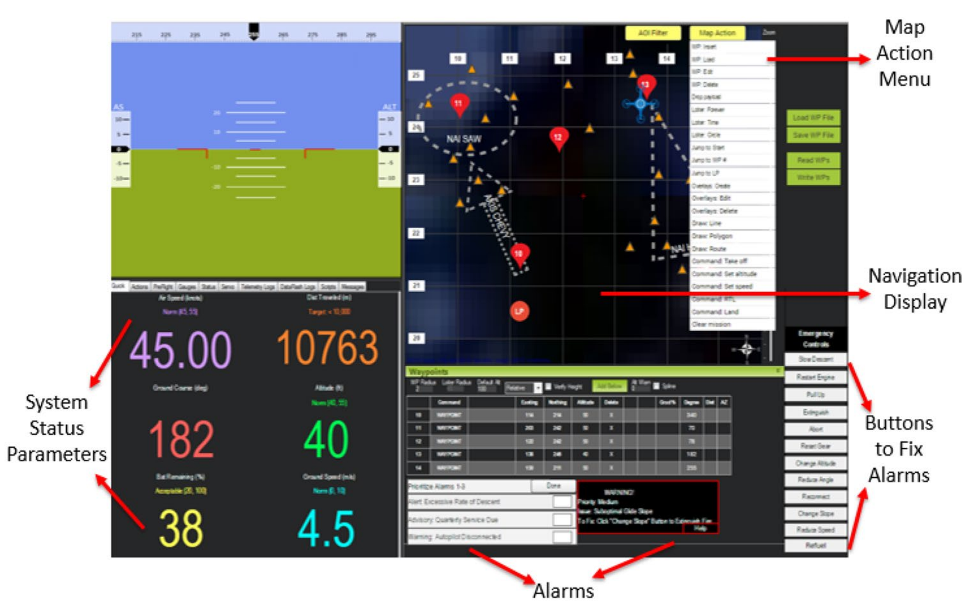

Fig. 2 Baseline interface prototype

Fig. 3 Enhanced interface prototype

improve operator understanding of the mission or system. For example, an object "distance tool" was made available and substantially simplified the distance estimation task by removing the need for users to visually measure distances relative to auxiliary guidelines. As another example of cueing, when the simulated UAV passed a waypoint, the color of the icon was changed to help the operator maintain awareness of mission progress.

Opposite to the enhanced interface, the degraded interface (Fig. 4) removed some task-related information from the interface, reduced information context and visual cueing for operators (see Table 5). The absence of these features did not prevent operators from being able to complete required tasks. However, the control tasks became more complex and challenging. For example, while the baseline interface provided some assistance in identifying resolutions to alarms or errors, the degraded prototype did not include these features. This situation caused operators to rely on memory of training and/or focus on the mission description document external to the interface (increasing task performance time).

2.4 Independent Variables

The experiment involved manipulation of two independent variables (IVs), including the interface design variation (i.e., baseline, enhanced, degraded) and the speed of the simulated UAV. The vehicle speed was set to either fast (a 5.5 min trial) or slow (an 8.5 min trial) and was maintained constant throughout testing scenarios. All trials included the same type and number of UAV control tasks and DKQs; therefore, the speed manipulation dictated the event rate for trials and represented a pure workload demand manipulation. As a result, participants had some idle time between tasks and/or DKQs under the slow condition but little idle time in the fast condition. The speed variable was meant to induce time pressure for participants under the fast conditions, thereby magnifying the demands of the control tasks, particularly in use of the degraded interface alternative. The speed manipulation also allows for assessment of the utility of the enhanced interface design (following transparency principles) for mitigating operator workload at a particular vehicle speed setting, in comparison to the other interface designs. The duration of a trial under the fast or slow condition was based on pilot testing by the authors. The fast trial duration (~ 5.5 min) was selected so that a participant had sufficient time for completing all tasks in the scenario, but with little idle time between tasks. The slow trial duration was set to approximately 50% longer than the fast trial.

2.5 Experiment Design

The experiment followed a mixed-factor design. The three interface settings served as a between-subject factor, while the speed condition was manipulated within-subjects. Each participant was randomly assigned to a single interface prototype in order to prevent user confusion among designs as well as the potential for learning or carryover effects when using multiple interfaces. Test trials followed one of two mission maps with the UAV flying at either the slow of fast speed for a total of four trials completed by each participant with their assigned interface. The maps served as replications of experiment condition (i.e., combinations of interface setting and speed). The geographical

Table 4 Differences between the Baseline and Enhanced Interface

Manipulation type	Interface component	Baseline interface	Enhanced interface	Related activity
Pre-attentive cueing	Alarms	Messages indicate the cause of alarms, provide basic information on priority and suggest solution	Messages indicate cause of alarms, provide priority information by color-coding, and highlight suggested solution with color and icon	Fix alarm; Prioritize alarm
Pre-attentive cueing	Areas of Interest (AOI)	A shortcut button provides a line coded filter (i.e. dotted margin) to identify AOIs	A shortcut button provides a shade coded filter (i.e. shaded area) to identify AOIs	Count targets; Determine target coordinates
Pre-attentive cueing and automation assistance	Flight Parameter Display	Flight parameters are displayed along with acceptable parameter ranges	Flight parameters are displayed along with acceptable ranges. Alert icons appear when parameters exceed acceptable ranges	Monitor flight parameter
Automation assistance	Target Presentation and Interaction	Guidelines and major coordinate markings present on navigation display. No direct target interaction through interface	Guidelines and major coordinate markings present. “Coordinate” and “Distance” shortcuts available to assist in user tasks	Determine coordinates; Estimate distance
Information context	Menu	Menu items are presented in a single list, grouped together by function	Menu items are presented in a hierarchical sub-menu structure per function	Locate menu options
Pre-attentive cueing	Waypoints	Waypoints are numbered. Traditional waypoint-style icons used to stand out from display background	Waypoints are numbered. Traditional waypoint-style icons used to stand out from display background. Waypoint color changed after the UAV flies past	Determine mission completion status

Fig. 4 Degraded interface design

features presented in each map were similar but different so as to prevent potential mission learning effects. The speed conditions and maps were both within-subject factors and were randomized for each participant.

2.6 Dependent Variables

2.6.1 Task Performance

Dependent measures included participant task performance, knowledge accuracy and workload responses. Performance measures included individual task completion times and task accuracy (or deviation) from the correct outcome. Both measures were analyzed for each of the previously identified control tasks. All participant responses were collected via screen capture and audio recording software applied during the test trials.

Coordinates Identification Task. The time for this task began with the first syllable of an audio command to a participant ("Report the coordinates…") and ended on the final syllable spoken by the participant in their response. Response deviation was measured in absolute Euclidian distance from the reported coordinates to the actual target coordinates. If a participant gave no response, the response deviation was not computable. Consequently, the trial was recorded as a "miss" and not included in data analysis.

Distance Estimation Task. The time for this task began with the first syllable of the audio command, "Report the distance…" and ended on the final syllable of the participant's response. Accuracy in this task was measured by the percent deviation of the reported distance from the actual (known) distance between two specified objects. For example, if the actual distance between two identified objects was 2500 m and the participant's response was 2000 m, then the deviation was calculated at 20% (= (2500–2000)/2500). If the participant gave no response, the distance estimation deviation was not computable. Consequently, a miss was recorded for the task and the data was not included in analysis.

Fix Alarm Task. The Fix Alarm task time began when an alarm appeared on the interface and ended on when the participant clicked the appropriate action button to resolve the alarm. In general, it was observed that participants tended to sacrifice task completion time for higher accuracy. All participants were able to correctly fix the alarm. Therefore, accuracy in this task was not analyzed.

Prioritize Alarm Task. The Prioritize Alarm task time began when alarms appeared on the interface and ended when a participant confirmed priority levels by clicking a "Done" button. Similar to the fix alarm task, most participants were able to correctly prioritize the alarms. Therefore, the accuracy for this task was not analyzed.

System Parameter Warnings. The System Parameter Warning task time began when a parameter value changed to an unacceptable number and ended when a participant finished reporting the abnormality to an experimenter. If no response was given, a miss was recorded and the observation was excluded from analysis.

2.6.2 Subjective Workload Ratings (NASA-TLX)

The NASA-TLX (Task Load indeX) was used as a subjective measure of participant cognitive workload. Hart and Staveland's [37] standard TLX forms and definitions were

Table 5 Differences between the Degraded and Baseline Interface

Manipulation	Interface component	Degraded	Baseline	Related activity
Lack of automation assistance	Alarms	Alarms only indicate a failure state and do not indicate cause. No priority information	Messages indicate the cause of alarms. Provides basic information on priority	Fix alarm; Prioritize alarm
Lack of context information	Areas of Interest (AOI)	No display filtering available	A shortcut button provides a line coded filter (i.e. dotted margin) to identify AOIs	Count targets; Determine target coordinates
No information context	Flight Parameter Display	Only the flight parameter is displayed	Flight parameters are displayed along with acceptable parameter ranges	Monitor flight parameter
No information context	Menu	Menu items are presented in a single list in random order	Menu items are presented in a single list, grouped together by function	Locate menu options
Lack of cueing	Target Presentation and Interaction	No guidelines on navigation display. Users cannot interact with targets in the navigation display through the interface	Guidelines and major coordinate markings present on navigation display. No direct target interaction through interface	Determine coordinates; Estimate distance
Lack of cueing and critical information	Waypoints	Colored similarly to navigation display background. Waypoints not numbered	Waypoints are numbered. Traditional waypoint-style icons used to stand out from display background	Determine mission completion status

used in this experiment. The 15 pair-wise comparisons of demand components were completed immediately following training trials. Each demand component was then rated immediately following each testing trial, for a total of four ratings of each demand component. Ratings were made on 5-inch bipolar visual analog scales with “low” and “high” anchors. Each rating was measured from the “low” anchor with a resolution of 1/16 of an inch and then transformed to a 100-point scale. The demand component rankings and ratings were combined to create an overall workload score for each test trial. The TLX score accounted for the ratings of physical, mental, temporal, effort, performance, and frustration demands.

2.6.3 DKQ Accuracy

The DKQs were used to measure an operator’s understanding of the UAV operating environment and system status. A total of 12 DKQs were posed during each test trial. The response accuracy was calculated as the percentage of correct answers over all query responses. If the participant did not respond to a DKQ, a miss was recorded and the related query was excluded from the accuracy calculation. (See Table 6 for DKQ misses and misses for all other experiment response measures.)

2.7 Procedure

2.7.1 Training

Participants were trained on the UAV control interface that they were assigned to. A researcher described each interface component, the location and function of features, and demonstrated use for tasks required during experiment trials. Participants were also familiarized with the mission scenario document, including trajectory map and the scheme of maneuver, as well as all types of DKQs that they could encounter during the experiment. Subsequently, participants completed a simplified mission as a training trial. The training trial mission presented all tasks described in Sect. 2.2, including the coordinates identification task, distance estimation task, alarm fix task, alarm prioritization task, and the parameter warning detection task. However, the mission map in the training trial was less geographically complex (see Fig. 5) and required a shorter duration to process (about 3 min), as compared to the test trials. In the case of automation assistance provided through the interface (e.g., alarm prioritization, distance tool), participants were told to trust the interface output as it would not provide incorrect suggestions. Similar to test trials, participants were also presented with auditory commands and 6 different DKQs during the training trial. An experimenter guided participant through all necessary actions to complete the training. The training

Table 6 Number of misses for dependent variables

Dependent variable	Enhanced interface	Baseline interface	Degraded interface
Coordinates identification (response deviation and time)	7	8	11
Distance estimation (response deviation and time)	0	0	5
Alarm fix time	0	0	3
Alarm prioritization time	0	0	3
System warning detection (rate and time)	0	4	3
NASA-TLX	0	0	0
DKQ accuracy	0	0	3
Total across all tasks	7	12	28

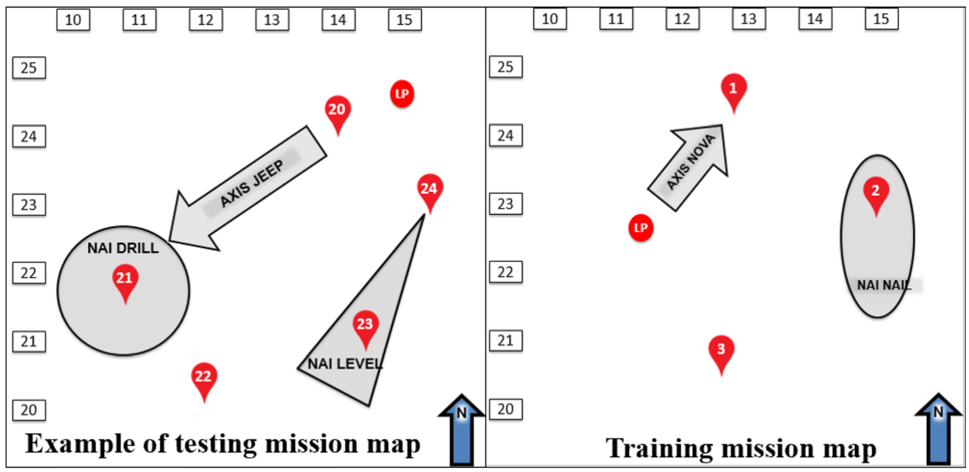

Fig. 5 Testing trial mission map (left) and training trial mission map (right)

mission was repeated until a participant was able to successfully address all tasks, answer all DKQs correctly, and report all system parameter warnings. Most participants repeated the training trial once. Before moving to test trials, participants ranked the TLX demand components based on the training mission scenario. In total, interface introduction, mission familiarization, and the training trials took approximately 45 min for each participant.

2.7.2 Test Trials

For a test trial, the mission included the same type of tasks. At the beginning of each test trial, participants were provided with an updated mission scenario document and time to review the scheme of maneuver. During test trials, auditory commands were presented via a digital message playback device and all DKQ answers and callouts of system parameter warnings were recorded by an experimenter. The UAV flew a predetermined route with set flight parameters. Immediately following each test trial, participants rated TLX demand components. Between trials, they were given a 2-min rest period. At the end of the experiment, they were paid for their time and dismissed.

2.8 Research Hypotheses

In general, the interface prototype manipulations were expected to have an impact on participant accuracy and task completion time during the simulated UAV missions. Specifically, the enhanced interface design was expected to yield superior performance to the baseline and degraded transparency designs, particularly under high task demands; i.e., the fast vehicle speed. It was hypothesized (H) that the enhanced interface would produce the greatest task accuracy (H1), shortest task completion time (H2), greatest DKQ response accuracy (H3), and lowest operator perceived cognitive load (H4), followed by the baseline interface, and then the degraded interface. It was also expected that any differences among the interface designs in terms of performance, system knowledge, and workload would be greater under the higher vehicle speed or event rate condition as compared to low speed.

2.9 Data Analysis

All DVs were analyzed with a split-plot analysis of variance (ANOVA): interface design was the whole-plot factor; vehicle speed was the split-plot factor. Participant nested within interface design was used as the whole-plot error term. Homoscedasticity and normality assumptions of the ANOVA were assessed with Barlett's and Shapiro–Wilk's tests, respectively. If there was any violation, transformations were applied to response measures (e.g., log or square root transformations) or the ranks of responses were submitted to the ANOVA procedure to perform a non-parametric analysis. However, when parametric and nonparametric procedures yielded identical results, the regular parametric analysis on the non-transformed responses was considered valid and reported [38]. A significance level of 0.05 was used as the criterion for establishing statistical significance of all ANOVA results. For main effects and interactions found to be significant, Tukey's Honest Significant Difference (HSD) test was used for pairwise comparison among treatment levels with the same significance criterion.

3 Results

The main effect of interface design variation was significant for all DVs. Test statistics are summarized in Table 4. Note that the denominator degree of freedoms vary due to missing data points with certain responses. Tukey's HSD post-hoc test results are indicated by letters in the plots of response measures shown in Fig. 5. Conditions labeled with the same letter were not significantly different from each other (Table 7).

The main effect of vehicle speed was only found to be significant for alarm fix time ($F(1, 44) = 4.27$, $p = 0.0446$), distance estimation task time ($F(1, 44) = 6.32$, $p = 0.0157$) and NASA-TLX scores ($F(1,45) = 16.73$, $p = 0.0002$). None of the DVs revealed significance of an interface design variation by vehicle speed interaction effect. In addition, a significant effect of trial number was observed for several of the DVs, including system warning detection rate ($F(1,44) = 4.42$, $p = 0.0412$), alarm fix time ($F(1,44) = 4.99$, $p = 0.0306$), alarm prioritization time ($F(1,44) = 5.41$ $p = 0.0247$) and TLX scores ($F(1,45) = 26.81$, $p < 0.0001$). The system warning detection rate tended to increase, and the task completion times tended to decrease in later trials. Despite the extensive mission training, these effects suggested that some participant learning may still have occurred during the course of the test trials (Fig. 6).

Table 7 ANOVA results on interface variation effect

DV	F-statistic	*P* Value
Coordinates identification task deviation	F(2,44) = 13.14	< .0001
Coordinates identification task time	F(2,44) = 13.94	< .0001
Distance estimation task deviation percentage	F(2,44) = 23.53	< .0001
Distance estimation task time	F(2,44) = 7.36	0.0017
Alarm Fix time	F(2,44) = 19.69	< .0001
Alarm prioritization time	F(2,44) = 26.92	< .0001
System warning detection rate	F(2,44) = 7.91	0.0012
System warning detection time	F(2,43) = 22.71	< .0001
NASA-TLX	F(2,45) = 3.24	0.0486
DKQ Accuracy	F(2,45) = 22.60	< .0001

4 Discussion

We hypothesized that the enhanced interface would generate superior task performance, followed by the baseline interface and then the degraded interface. This hypothesis was partially supported by the results. For the warning detection and coordinates identification tasks, both the enhanced and baseline interfaces produced greater task performance than the degraded interface. However, no there was no significant difference between the enhanced and baseline interfaces. The lack of cueing and information context in the degraded interface had a detrimental effect on task performance. However, further improvement of cueing features and the addition of automation assistance provided by the enhanced interface did not significantly improve performance in these two tasks relative to the baseline condition. With respect to the distance estimation task, significant differences among the interface conditions were in line with expectation. The enhanced interface automation assisted participants with the cognitive process of determining the distance between objects or exact coordinates of objects, leading to significantly improved accuracy as compared to the baseline interface. (Anecdotally, these were the most difficult tasks according to participants.) In regard to participant accuracy in addressing alarms, there was a negligible difference between the interface designs. Despite the automation advantage of the enhanced interface users, task response accuracies were indistinguishable and were all at 100%. Participants were instructed to execute each task as quickly and accurately as possible; however, the experiment results made clear that participants placed greater emphasis on accuracy. For this reason, the accuracy responses for these two tasks were not analyzed.

The second hypothesis stated that participants using the enhanced interface would produce faster task completion times. This hypothesis was also partially supported by the results. For the distance estimation and alarm prioritization tasks, the trend of significant differences among the three interfaces was as expected. However, the difference in task completion times was not significant between the baseline and degraded interfaces for the parameter warning detection, fix alarm, and coordinates identification tasks. These results may be, in part, attributable to the enhanced interface providing some automation of tasks that was not available to

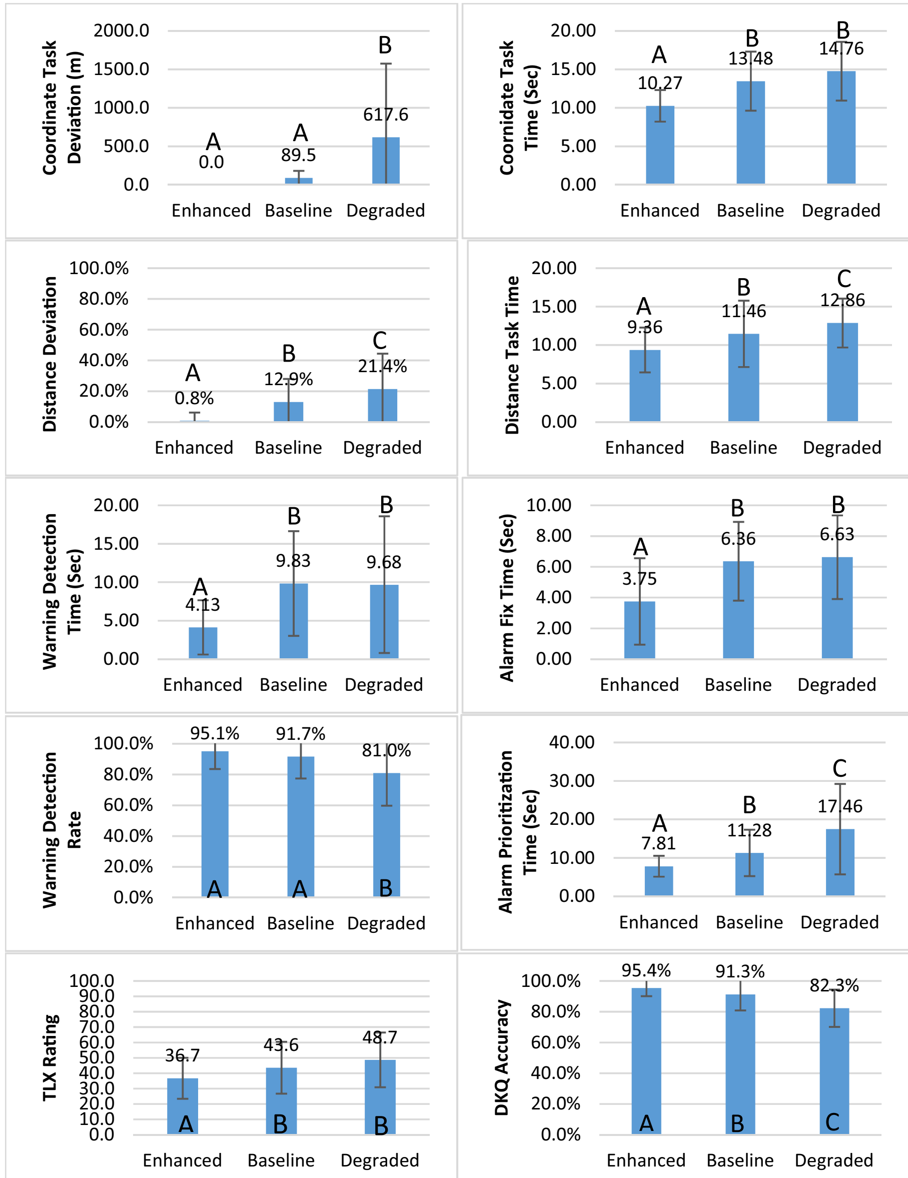

Fig. 6 Plots of response measure means (and standard deviations) for interface design conditions along with Tukey's HSD results

other users. More specifically, the coordinates identification and distance estimation task tools both represented information analysis automation for the cognitive processes of object location coordinates and inter-object distance estimation, respectively. Related to the fix alarm task, the enhanced interface provided pre-attentive cueing features by highlighting the alarm solution suggested by the system. This feature eliminated the need for participant visual search of a long list of possible actions and, therefore, significantly reduced the task time. Similarly, the color-coding of alarms suggested priorities for completion of the prioritize alarm task. This feature removed the need for reading alarm text and reduced task performance time. Regarding the system parameter warnings task, automation assistance was provided so that the abnormal status of system parameters could be easily recognized by users of the enhanced interface. This eliminated the need for visual search and mental analysis.

The third hypothesis posited that the enhanced interface would produce superior DKQ accuracy, followed by the baseline interface and then the degraded interface. Experiment results supported this hypothesis. By providing pre-attentive cueing, the enhanced interface successfully improved user awareness of mission status. The enhanced interface also made the map action menu more usable by implementing a hierarchical structure and logical grouping of options. The degraded interface, on the other hand, removed useful cues (e.g., waypoints were not easily distinguishable from the map background) and task-critical information (e.g., AOI filters or identifiers). The absence of these features caused participants to frequently refer to printed mission materials and to rely on working memory for task performance. Consequently, operator capability to maintain an accurate understanding of the UAV environment and subsystem status was reduced.

However, one limitation of the DKQs is that they were not designed to capture each level of SA-based transparency. When examined retrospectively, we found that the majority of DKQs addressed Level-1 transparency, concerning basic information on the agent, environment, or the mission. Only a few DKQs were relevant to the "reasoning" aspect of Level-2 transparency (e.g., How many targets are inside [AOI name]?) or the future projection aspect of Level-3 transparency (e.g., What is your current completion percentage for this mission?). While acknowledging this limitation, we think the impact of the interface might be inferred from performance of the alarm prioritization, coordinates identification, and distance estimation tasks, which would not be effective without interface support beyond Level-1 transparency.

It was also hypothesized (H4) that operator workload would be the lowest in using the enhanced interface, followed by the baseline interface and then the degraded interface. Experiment data partially supported this hypothesis. Participants using the enhanced interface produced significantly lower TLX scores than both the baseline and degraded interface users. However, there was no significant difference between the baseline and the degraded interface groups. The reason might be that more automation assistance features were available in the enhanced interface. Prior human-automation interaction research [39–41] has demonstrated substantial performance and workload benefits when automation is applied in a manner relative to system user information processing demands.

While missed responses for the dependent variables were not included in the data analysis, it was observed that the enhanced interface resulted in fewere misses and the degraded interface resulted in more misses. This observation further demonstrates the effects of the interface manipulations.

In this experiment, we also attempted to generate high task demands by manipulating UAV flight speed and, consequently, mission task rate. This manipulation was only successful for certain responses, including alarm fix time, distance estimation task time and NASA-TLX scores. That is, task time and workload increased with the higher speed setting. However, it is possible that the high vehicle speed was not sufficiently demanding to affect participant performance in the other common control tasks. An additional intent of the speed manipulation was to determine if the enhanced interface would be more robust for supporting operator performance under high demand circumstances. However, the present results did not allow for investigation of any such relationship.

The significant trial number effect indicated that additional task learning may have occurred during experimental trials. Despite extensive training (about 45 min), an improvement in task performance might have occurred due to additional proficiency, more participant confidence and increased psychological comfort over the course of the experiment.

In general, our findings were consistent with other studies in revealing benefits of automation transparency features. For example, Mercado [22] investigated the effects of level of agent transparency in the context of human–agent teaming for multi-robot management. Results indicated that operator performance, trust, and perceived usability increased as a function of transparency level. Subjective and objective workload data indicated that participant workload did not increase, as a function of transparency. Wright [21] also examined how the transparency of an agent's reasoning affected human operator tendency for complacency in a military route selection task. It was found that improved transparency through access to agent reasoning effectively improved human performance and decreased automation bias.

Unlike Mercado [22], the present study did not align the interface manipulations with the three levels of agent transparency defined by Chen et al. [13]. We applied or violated as many transparency design guidelines in the literature as feasible to improve or degrade transparency of the interface and agent in the target application. However, taking a retrospective view, we found some association between the interface manipulations in the present study and the three levels of transparency defined by Chen et al. [13]. Pre-attentive cueing in the enhanced interface mainly improved Level-1 transparency by making information on the agent, environment, or task easier to process. For example, shading of the AOIs on the map made it easy to process information related to the environment as well as the task. Manipulations on information context mainly impacted Level-2 transparency by changing how well environmental/vehicular constraints were represented in the interface. For example, the lack of normal parameter range information in the degraded interface required operators to memorize them from the mission description and, therefore, impaired "constraint-driven reasoning" when determining the status of the vehicle. The automation assistance in this study was the only type of manipulation that impacted Level-3 transparency. For example, the alarm priority and recommended actions were no longer automatically determined in the degraded interface. This made operator projection of future actions and system states more difficult. While the distance estimation and coordinates tool in the enhanced interface were not directly relevant to future projections of the task or environment in the experiment, such automation assistance could be critical for Level-3 transparency in other scenarios (e.g., the mission requires an unexpected route change).

In addition, the present study differed from prior automation transparency research in several ways. First, the automation features presented in the enhanced interface condition were "perfect". That is, there were no reliability issues in the function of the automated aids. Participants were instructed to completely trust the automation. In this way, we observed a "pure" effect of transparency efforts without confounding of automation reliability issues, as in some previous studies. On the other hand, the "perfect" automation assumed in this study (by necessity) decreased decision making time and cognitive workload in evaluating the reliability of automation recommendations. This special constraint makes the research results applicable to situations where automation is highly reliable and user trust is not a challenge. Secondly, many interface usability features that were not previously identified as "transparency" measures were explored in the design variations of this research. Various response measures revealed these features to be effective for building a transparent, automated system that was helpful for promoting operator performance and understanding of mission status. Based on the results of this study, we receommend further exploration of usability guidelines in UAV interface design to support automation transparency. Finally, while most prior research has been aimed at achieving transparency benefits at no cost of workload, the present study showed that an enhanced transparency interface can actually reduce user workload through careful identification of automated functionality and usability features.

5 Conclusion

5.1 Summary

In this research, we investigated the utility of functional transparency and usability features in UAV supervisory control interfaces. Following design recommendations from previous automation transparency research, we created three interface design variations: a baseline interface mimicking a commercially available UAV control station, an enhanced interface designed for increased automation transparency and usability, and a degraded interface removing some important task-related features. Results showed that the enhanced interface improved operator task performance, dynamic knowledge of UAV states, and reduced user workload. Additionally, the degraded interface showed detrimental effects of the absence of transparency and usability features. Among the interface features explored in this study, it appeared that automation assistance primarily contributed to task completion times and mitigated workload. However, the availability of pre-attentive cueing and information context was also important for accurate task performance.

By defining and setting automation transparency as an overarching design objective, we were able to prototype an enhanced interface that more effectively supported human-automation interaction in simulation of a realistic UAV control operation. Automation transparency as a high-level design objective may be useful for expert designers with understanding of human performance consequences of "clumsy" automation; whereas, usability design guidelines, as "building blocks" to transparency, may be useful for new system designers.

5.2 Limitations and Future Work

One limitation of this study was the reduced interactivity of the interface prototypes presented to participants. Functional features were limited to those with direct relevance to performance of the required UAV control tasks. Furthermore, if a user made erroneous selection of an inactive function, there was no negative impact on vehicle function. In the real-world, misuse of UAV control interface features can lead to mission failure or the loss of a vehicle. In order to focus our study on the impact of automation transparency on specific UAV task performance, the possibility for total failure was removed from the simulation.

The goal of this research was to investigate the effect of interface transparency on task performance. While the study carefully evaluated performance measures and the dynamic knowledge of participants, additional validation on the level of transparency would have been beneficial. For example, following the application of transparency design guidelines to an interface, a study could be conducted to evaluate the extent to which different levels of situation awareness are achieved or impaired by the interfaces with different levels of transparency. Such a study would also allow us to further delineate how additional functionality and usability features drive the level of transparency and, consequently, situation awareness and performance.

Beyond this, the study only made use of three interface designs, which were all based on the same commercially available GCS; the ArduPilot MP. However, at present, there are a substantial number of GCSs available on the market and many more military-grade GCSs.

While this study aimed to investigate performance in common UAV control operations, future work should explore additional GCS design variations with different automated features that may further improve operator performance and reduce cognitive workload. Future studies may also examine the use of such interfaces under off-nominal UAV flight conditions as well as in situations where recommended actions by automation are not always reliable. Such research could further the understanding of how automation transparency, through control interface design, might serve to also support operator performance in knowledge-based tasks under high demand circumstances. Lastly, the current research only recruited participants with no disability and/or color impairment. Future research should be conducted to explore UAV control interface design following inclusive design principles [14]. Such an effort is critical for enabling a broader population of users in UAV control operations.

Funding This research was supported by a grant from the National Aeronautics and Space Administration (NASA Grant No. NNX16AB23A). Terry Fong was the technical monitor. The views and opinions expressed are those of the authors and do not necessarily reflect the views of the NASA.

Compliance with Ethical Standards

Conflict of interest The authors declare that they have no conflict of interest.

Research involving human participants The institutional review board of North Carolina State University has approved this study. All participants were provided information on the objectives, procedure and potential risks related to the experiment and signed an informed consent form prior to testing.

References

1. Nullmeyer R, Herz R, Montijo G (2009) Training interventions to reduce air force predator mishaps. Air Force Research Lab Mesa AZ Human Effectiveness Directorate. Report No.: AFRL-RH-AZ-PR-2009–0002. Available from: http://www.dtic.mil/docs/citations/ADA501747
2. Tvaryanas AP, Thompson WT, Constable SH (2006) Human factors in remotely piloted aircraft operations: HFACS analysis of 221 mishaps over 10 years. Aviat Space Environ Med 77(7):724–732
3. Yesilbas V, Cotter TS (2014) Structural analysis of Hfacs in unmanned and manned air vehicles. In: Proceedings of the international annual conference of the American Society for Engineering Management; Huntsville. Huntsville: American Society for Engineering Management (ASEM). pp 1–7. Available from: http://search.proquest.com/docview/1705171380/abstract/5BCAB252E0054797PQ/1
4. Giese S, Carr D, Chahl J (2013) Implications for unmanned systems research of military UAV mishap statistics. In: 2013 IEEE intelligent vehicles symposium (IV), pp 1191–6.
5. Hobbs A (2010) Unmanned aircraft systems. Hum Factors Aviat, pp 505–531
6. Fong T, Thorpe C (2001) Vehicle teleoperation interfaces. Auton Robots 11(1):9–18
7. Williams KW (2006) Human factors implications of unmanned aircraft accidents: flight-control problems. In: Human factors of remotely operated vehicles. Emerald Group Publishing Limited, pp 105–116. Available from: http://www.emeraldinsight.com, https://doi.org/10.1016/S1479-3601(05)07008-6
8. Neville K, Blickensderfer B, Archer J, Kaste K, Luxion SP (2012) A cognitive work analysis to identify human-machine interface design challenges unique to uninhabited aircraft systems. Proc Hum Factors Ergon Soc Annu Meet 56(1):418–422
9. Hobbs A, Lyall B (2016) Human factors guidelines for unmanned aircraft systems. Ergon Des 24(3):23–28
10. Pedersen HK, Cooke NJ, Pringle H, Connor O (2006) UAV human factors: operator perspectives. In: Cooke NJ, Pringle H, Pedersen H, Connor O (eds) Human factors of remotely operated vehicles
11. Sarter NB, Woods DD (1995) How in the world did we get into that mode? Mode error and awareness in supervisory control. Hum Factors 37:5–19
12. Selkowitz AR, Larios CA, Lakhmani SG, Chen JYC (2017) Displaying information to support transparency for autonomous platforms. In: Advances in human factors in robots and unmanned systems. Springer, Cham, pp 161–73. (Advances in Intelligent Systems and Computing). Available from: https://link.springer.com/chapter, https://doi.org/10.1007/978-3-319-41959-6_14
13. Chen JY, Procci K, Boyce M, Wright J, Garcia A, Barnes M (2014) Situation awareness-based agent transparency (No. ARL-TR-6905). Army research lab Aberdeen proving ground MD human research and engineering directorate.
14. Lyons JB (2013) Being transparent about transparency: a model for human-robot interaction. In: 2013 AAAI spring symposium series [Internet]. Available from: https://www.aaai.org/ocs/index.php/SSS/SSS13/paper/view/5712
15. Degani A, Shafto M, Kirlik A (1999) Modes in human-machine systems: constructs, representation, and classification. Int J Aviat Psychol. 9(2):125–138
16. DeVisser EJ, Cohen M, Freedy A, Parasuraman R (2014) A design methodology for trust cue calibration in cognitive agents. In: Virtual, augmented and mixed reality designing and developing virtual and augmented environments. Springer, Cham, pp 251–62. (Lecture Notes in Computer Science). Available from: https://link.springer.com/chapter, https://doi.org/10.1007/978-3-319-07458-0_24
17. Masalonis AJ, Parasuraman R (2003) Effects of situation-specific reliability on trust and usage of automated air traffic control decision aids. Proc Hum Factors Ergon Soc Annu Meet 47(3):533–537
18. Kilgore R, Voshell M (2014) Increasing the transparency of unmanned systems: applications of ecological interface design. In: Virtual, augmented and mixed reality applications of virtual and augmented reality. Springer, Cham, pp 378–89. (Lecture Notes in Computer Science). Available from: https://link.springer.com/chapter, https://doi.org/10.1007/978-3-319-07464-1_35
19. Selkowitz A, Lakhmani S, Chen JYC, Boyce M (2015) The effects of agent transparency on human interaction with an autonomous robotic agent. Proc Hum Factors Ergon Soc Annu Meet. 59(1):806–810
20. Selkowitz AR, Lakhmani SG, Larios CN, Chen JYC (2016) Agent transparency and the autonomous squad member. Proc Hum Factors Ergon Soc Annu Meet. 60(1):1319–1323
21. Wright JL, Chen JYC, Barnes MJ, Hancock PA (2016) The effect of agent reasoning transparency on automation bias: an analysis of response performance. In: Virtual, augmented and mixed reality. Springer, Cham, pp 465–77. (Lecture Notes in Computer Science). Available from: https://link.springer.com/chapter, https://doi.org/10.1007/978-3-319-39907-2_45
22. Mercado JE, Rupp MA, Chen JYC, Barnes MJ, Barber D, Procci K (2016) Intelligent agent transparency in human-agent teaming for multi-UxV management. Hum Factors 58(3):401–415
23. Sadler G, Battiste H, Ho N, Hoffmann L, Johnson W, Shively R (2016) Effects of transparency on pilot trust and agreement in the autonomous constrained flight planner. In: 2016 IEEE/AIAA 35th digital avionics systems conference (DASC). pp 1–9
24. Dixon SR, Wickens CD (2003). Control of multiple-UAVs: a workload analysis. DTIC Document.

25. Galster SM, Knott BA, Brown RD (2006). Managing multiple UAVs: are we asking the right questions? In: Proceedings of the human factors and ergonomics society annual meeting, vol 50, pp 545–549
26. Squire P, Trafton G, Parasuraman R (2006) Human control of multiple unmanned vehicles: effects of interface type on execution and task switching times. In: Proceedings of 1st ACM SIGCHI/SIGART conference on Human-robot interaction. ACM, pp 26–32
27. Prewett MS, Johnson RC, Saboe KN, Elliott LR, Coovert MD (2010) Managing workload in human–robot interaction: a review of empirical studies. Comput Hum Behav 26(5):840–856
28. Peschel JM, Murphy RR (2013) On the human-machine interaction of unmanned aerial system mission specialists. IEEE Trans Human-Mach Syst 43(1).
29. Afergan D, Peck EM, Solovey ET, Jenkins A, Hincks SW, Brown ET, Jacob RT (2014) Dynamic difficulty using brain metrics of workload. In: Proceedings of the 32nd annual ACM conference on human factors in computing systems. ACM, pp 26–32
30. Chen JY, Barnes MJ, Harper-Sciarini M (2011) Supervisory control of multiple robots: human-performance issues and user-interface design. IEEE Trans Syst Man Cybern C 41(4):435–454
31. Cummings ML, Myers K, Scott SD (2006) Modified Cooper Harper evaluation tool for unmanned vehicle displays. In: Proceedings of UVS Canada: conference on unmanned vehicle systems Canada.
32. Fahlstrom P, Gleason T (2012) Introduction to UAV systems. Wiley
33. Fuchs C, Borst C, de Croon GC, van Paassen MR, Mulder M (2014) An ecological approach to the supervisory control of UAV swarms. Int J Micro Air Vehicles 6(4):211–229
34. Lorite S, Muñoz A, Tornero J, Ponsa P, Pastor E (2013) Supervisory control interface design for unmanned aerial vehicles through GEDIS-UAV. In: International conference on human-computer interaction, pp 231–240. Springer
35. Lu JL, Horng RY, Chao CJ (2013) Design and test of a situation-augmented display for an anmaned aerial vehicle monitoring task. Perceptual Motor Skills 117(1)
36. Menda J, Hing JT, Ayaz H, Shewokis PA, Izzetoglu K, Onaral B, Oh P (2011) Optical brain imaging to enhance UAV operator training, evaluation, and interface development. J Intell Rob Syst 61(1–4):423–443
37. Hart SG, Staveland LE (1988) Development of NASA-TLX (task load index): results of empirical and theoretical research. Adv Psychol 52: 139–83. (Human Mental Workload; vol. 52). Available from: http://www.sciencedirect.com/science/article/pii/S0166411508623869
38. Conover WJ, Conover WJ (1980) Practical nonparametric statistics. Available from: https://orion.sfasu.edu/courseinformation/syl/201202/MTH4752.pdf
39. Parasuraman R, Sheridan TB, Wickens CD (2000) A model for types and levels of human interaction with automation. IEEE Trans Syst Man Cybern Part Syst Hum 30(3):286–297
40. Kaber DB, Wright MC (2003) Adaptive automation of stages of information processing and the relation to operator functional states. NATO Sci Ser SUB Ser Life Behav Sci 355:204–223
41. Kaber DB, Endsley MR (2004) The effects of level of automation and adaptive automation on human performance, situation awareness and workload in a dynamic control task. Theor Issues Ergon Sci. 5(2):113–153

Wenjuan Zhang is currently a Research Scientist on Human-Computer Interaction at Dataminr Inc. She received her Ph.D. in Industrial Engineering at North Carolina State University in 2017 and her Master of Science in the same program in 2015. Her research interests include human-automation interaction, system design, user interaction with emerging technologies.

David Feltner is currently a professor in the Engineering Psychology department at West Point. Feltner obtained his Master of Science in Industrial Engineering program at North Carolina State University in 2017 and his Bachelors of Science at the United States Military Academy in 2008.

David Kaber is the Department Chair and Dean's Leadership Professor in the Department of Industrial & Systems Engineering at the University of Florida. He is also an Adjunct Professor of Industrial and Systems Engineering at North Carolina State University. Kaber received his PhD from Texas Tech University in 1996. He is a Certified Human Factors Professional and Fellow of HFES and IIE. He is currently serving as Editor-in-Chief for the IEEE Transactions on Human-Machine Systems.

James Shirley is currently a Reliability Engineer at NAVAIR. He received his Bachelor of Science in Industrial Engineering degree in 2016 and his Master of Science in the same program in 2018.