Moving Autonomy Forward Conference 2006
De Vere Hotel Belton Woods
Lincoln, UK

HUMAN PERFORMANCE CONSIDERATIONS IN THE DEVELOPMENT OF INTEROPERABILITY STANDARDS FOR UAV INTERFACES

M.L. Cummings, J.T. Platts, A. Sulmistras
MIT, Muretex, CDL Systems

Human Performance Considerations in the Development of Interoperability Standards for UAV Interfaces

M.L. Cummings, J.T. Platts, A. Sulmistras

ABSTRACT

Against a background of parallel UAV development and procurement programs, NATO nations are faced with the potential likelihood of many similar systems both within their forces and between coalition countries. NATO STANAG 4586 was developed to decouple the control station from the air vehicle in UAV systems, thus presenting the option of developing a single control station to control a range of UAV types in disparate missions and also to allow for greater interoperability. This paper discusses the human supervisory control implications of such interoperability standards in the development of UAV ground control interfaces. Interoperability standards are not only critical for improving current NATO operations, but will also lay the groundwork for future network centric operations. However, to achieve greater flexibility and cost effectiveness with unmanned aerial systems (UAS), more system intelligence will be needed in order to improve tempo of information sharing, reduce operational incidences, and ensure overall mission effectiveness. Thus interoperability standards of the future must be able to support such functionalities.

INTRODUCTION

Network Centric Operations (NCO), also known as Network Centric Warfare (NCW), is a concept of operations envisioned to achieve information superiority and mission success by effectively linking or networking knowledgeable entities in a battlespace [2]. As defined by the U.S. Department of Defense, key components of NCO include information sharing and collaboration which will promote shared situational awareness and overall mission success. However, a significant challenge in the implementation of NCO will be achieving the rapid information sharing not only intra- but inter-nation as well. One particular area of networked operations where multinational NCO could be problematic is in the dissemination of information and images from UAVs. To promote true interoperability between multinational forces, all relevant stakeholders will need access to data generated by UAVs in a network but significant differences in hardware and software across a multinational force presents a major hurdle.

In an attempt to address the problem of interoperability between nodes in a multinational UAV network, the NATO Standardisation Agreement (STANAG) 4586 was created to address standard interfaces of UAV control systems. In its second edition, STANAG 4586 was conceptualized to promote interoperability between one or more ground stations, unmanned air vehicles and their payloads, and the Command, Control, Communication, Computer and Intelligence (C4I) network, particularly in joint operational settings. STANAG 4586 attempts to accomplish this through implementing "standard interfaces", which should not be confused with human-computer interfaces (HCI). The standard interfaces are essentially communication message sets between the vehicle and a ground control station. Figure 1 is a simplified representation of the

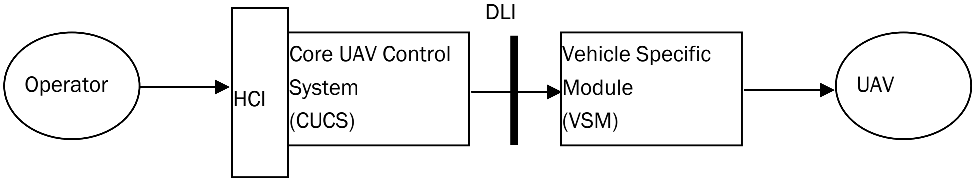

Figure 1: Basic STANAG 4586 Architecture [1]

STANAG 4586 architecture. An operator interfaces with a ground control station (known as the Core UAV Control System (CUCS)), which communicates through message sets in a data link interface (DLI) with the UAV, specifically through the Vehicle Specific Module (VSM), which may or may not be on the actual vehicle.

Five Levels of Interoperability (LOI) have been delineated for STANAG -compliant UAV systems:

- Level 1: Indirect receipt/transmission of UAV related payload data
- Level 2: Direct receipt of ISR/other data where "direct" covers reception of the UAV payload data by the UCS when it has direct communication with the UAV.
- Level 3: Control and monitoring of the UAV payload in addition to direct receipt of ISR/other data
- Level 4: Control and monitoring of the UAV, less launch and recovery
- Level 5: Control and monitoring of the UAV (Level 4), plus launch and recovery functions

The increasing LOIs represent increasing human interaction with a vehicle, but they do not necessarily imply a higher lever of autonomy. A LOI 5 vehicle could be a Predator which is hand-flown but it could also be a Shadow that an operator chooses to land via the automated recovery system. Thus the LOIs do not represent levels of autonomy, they only imply how much interaction/authority one or more human controllers could expect to have with a system. Within each LOI, there could be additional layers of autonomy. Because they represent the data elements that must be communicated between the ground control station and the vehicle (as seen by the operator through the interface), the message sets outlined in STANAG 4586 can inform ground control station HCI design.

Several vehicles have flown using STANAG software (CDL Systems Vehicle Control Station)[1], including Silver Fox (Advanced Ceramics Research), Grasshopper (Xiphos), Warrior/ERMP (General Atomics)[2], Kingfisher (BAE Systems), and Shadow and Pioneer (AAI) are to be flown within the year. The U.S. Navy has planned to implement STANAG 4586 through the Raytheon Tactical Control Station (TCS)[3] for Firescout, Dragon Eye (USMC), and Predator, funded by JFCOM as part of JOTBS.

HUMAN SUPERVISORY CONTROL AND UAV OPERATIONS

What it means to "control" a UAV from the human perspective is an issue that is central to the development of interoperability standards. Human control of UAVs is fundamentally a supervisory control task since operators are remote supervisors and managers of highly complex systems nested in the highly uncertain and dynamic environment of command and control. This remote, computer-mediated form of control is known a human supervisory control (HSC), which is the process by which a human operator intermittently interacts with a computer, receiving feedback from and providing commands to a controlled process or task environment, which is connected to that computer (Figure 2). This diagram is similar to the STANAG 4586 architecture (Figure 1).

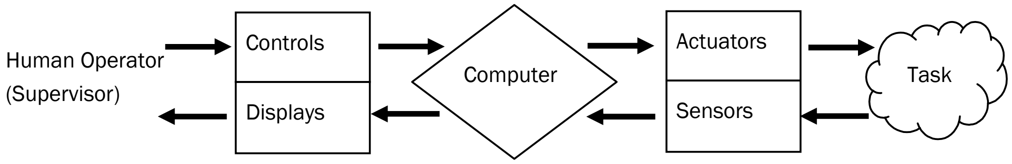

Figure 2: Human Supervisory Control

---

[1] The STANAG 4586 custodian, US Navy PMA-263, has not fully put in place validation and conformance testing, thus no company/vehicle can yet claim meeting certification or being tested to full compliance.

[2] In the United States, STANAG 4586 has been mandated as a requirement for a number of systems (e.g., BAMS, U.S. Army ER/MP) 109th Congress, H.R. 1815 Section 141. 2005, U.S. Congress

[3] Raytheon has a new TCS-derivative ground control station, the Multiple Vehicle Control System (MVCS), which is planned to be STANAG compliant.

Figure 3 represents the embedded control loops that occur in supervisory control of one or more UAVs. The inner loop of Figure 3 represents a basic guidance and motion control loop which is the most critical loop that must obey physical laws of nature such as aerodynamic constraints. The second loop, the navigation loop, represents the actions that some agent, whether human or computer-driven, must execute to meet local geo-spatial constraints such as routes to waypoints or points of interest, time on targets, and avoidance of obstacles. The navigation loop corresponds to the control of the path necessary to accomplish the mission. The outermost loop represents the highest levels of control, that of mission management. It is this loop where typically the actual mission requirements are met. For example, for UAV missions such as intelligence, surveillance, and reconnaissance (ISR), sensors must be monitored and decisions made based on the incoming information to meet overall mission requirements. Finally, the system health and status monitoring loop represents the continual supervision that must occur, either by a human or automation or both, to ensure all systems are operating within normal limits.

From the human-in-the-loop perspective, if the inner loops fail, then the higher or outer loops will also fail. The dependency of higher loop control on the successful control of the lower loops drives human limitations in control of one or more UAVs. If humans must interact in the guidance and motion control loop (i.e., hand-fly a UAV) the human cost is high because this effort requires significant cognitive resources. What little spare mental capacity is available must be divided between the navigation and mission management control loops.

Violations of the priority scheme represented in Figure 3 have led to serious problems exemplified by numerous Predator crashes. When operators become cognitively saturated or do not correctly allocate their cognitive resources to the appropriate control loops in the correct priorities, they violate the control loops constraints, potentially causing catastrophic failure. Moreover, in the health and status loop, operators could fixate on a system problem that may or may not be critical, and possibly cause the failure of an inner loop due to inappropriate attention allocation. The challenge in achieving effective management of either a single or multiple UAVs in the future is not only to determine if automation can be used to reduce workload, but to what degree in each of the control loops in Figure 3.

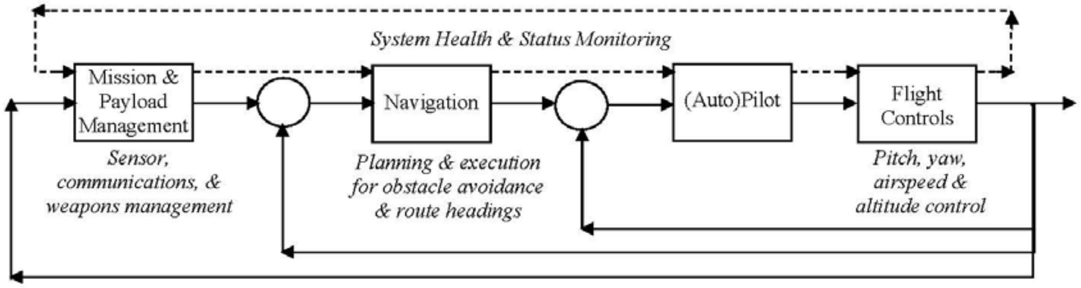

Figure 3: Nested Control Loops for Supervision of UAVs

Levels of Automation

The challenge in achieving efficient, accurate, and safe unmanned vehicle operations in the future is not only to determine what can be automated, but to what degree in each of the control loops in Figure 3. Automation strategies can range from fully automatic where the operator is completely left out of the decision and control process to minimal levels where the automation offers basic data filtering or recommendations for the human to consider. Sheridan and Verplank outlined a scale from 1-10 where each level represents progressively more automation for decision and action selection, as shown in Table 1 [3]. For lower-level perceptually-based tasks such as the inner loop in Figure 3, higher levels of automation (LOAs) in general result in lower workload, e.g. autopilot. However, it is much more difficult to automate tasks in the outer two loops, particularly the mission management loop in Figure 3, because they require knowledge-based decision making (e.g., experience, judgment, pattern recognition) that is extremely difficult to automate.

Table 1: Levels of Automation

Automation Level	Automation Description
1	The computer offers no assistance: human must take all decision and actions.
2	The computer offers a complete set of decision/action alternatives, or
3	narrows the selection down to a few, or
4	suggests one alternative, and
5	executes that suggestion if the human approves, or
6	allows the human a restricted time to veto before automatic execution, or
7	executes automatically, then necessarily informs humans, and
8	informs the human only if asked, or
9	informs the human only if it, the computer, decides to.
10	The computer decides everything and acts autonomously, ignoring the human.

The key to successful UAV operations, whether it is a team of operators controlling one vehicle or the reciprocal is the introduction of the proper automation levels (Table 1) across the three nested control loops in Figure 3. To move towards network-centric operations (including interoperability), higher levels of automation must be introduced at all levels.

HUMAN SUPERVISORY CONTROL AND INTEROPERABILITY INTERFACE STANDARDS

STANAG 4586 can be seen as a representation of what information requirements are needed to support HSC functionalities. As such, the message sets and detailed data elements contain implicit, and often explicit, ramifications for successful human system integration. An analysis of the STANAG 4586 message sets and their related data elements revealed six areas that deserve further discussion in terms of human supervisory control.

- Autonomy support
- Mode confusion
- Information overload
- Message/display integration
- Feedback
- Future distributed architectures

It is important to note that these six areas are not mutually exclusive, and they are also not problem areas specific to STANAG 4586 and can be found across a number of HSC domains such as manned aviation, process control plant supervision, and even in the medical community.

Autonomy support

STANAG 4586 was originally conceptualized as an interoperability standard to promote higher-level autonomous operations, meaning that operators would not be responsible for basic piloting tasks but would primarily focus their attention on mission and payload management. In order to allow participation of legacy vehicles that operated more on a remotely-piloted paradigm (thus a lower level of autonomy), STANAG 4586 was revised to include message sets that would allow operators lower level, manual control.

Thus, STANAG 4586 allows for both lower and higher-level, more autonomous control. While there is a debate whether or not this duality should exist in STANAG 4586, the fact remains that if true interoperability is going to exist, this standard must allow for inclusion of all levels of autonomy, and thus both message sets are needed. Furthermore, there is significant organizational resistance to fully autonomous vehicles (e.g., the U.S. Air Force insistence that a Predator be “flown” by a human.” This topic is currently under debate for the U.S. Federal Aviation Administration, which currently asserts that any UAV operating in the National Airspace

System (NAS) must be directly controlled by a human. Thus, interoperability not only involves NATO military operations, but also civil as well, thus for the foreseeable future, this duality of the message sets must exist.

In terms of levels of automation within the mission planning and health and status monitoring message sets, there are important areas that need to be addressed. One fundamental issue is whether a not a system will operate on a management-by-consent (MBC) level (for which the operator approves all actions), or on a management-by-exception (MBE) level (for which the automation gives the operator some period of time to reject the automation's decisions/actions.) MBE will be critical for monitoring and intervening with complex health and status systems, but may or may not be applicable to payload management. The MBC/MBE allocation will be a significant issue in the development of electronic checklists, but is also critical for message sets that apply to more autonomous vehicles. Particularly for systems operating under a MBE strategy, it is imperative that the message sets contain all the relevant data elements such as time constraints for human intervention, as well as all the correct feedback data elements. Moreover they must occur in a timely fashion such that an operator can make the decision in the period of allotted time, thus any time latency introduced should be considered. Developing standardized MBE message sets is possible, but not straightforward (see Section V for a more comprehensive automation discussion).

Mode confusion

In any system that is both highly automated but also allows operator intervention during various phases of operation, there is a significant risk of operator "mode confusion". As evidenced by numerous airline crashes in the 1990's during the transition to highly automated flight control and navigation systems, mode confusion occurs when an operator attempts to take control of a highly automated system, but does not understand the current mode. This lack of understanding causes system failure due the confusion over who is in control (system or operator) and what should be done to achieve some desired goal state.

An example of mode confusion is the fatal Airbus320 crash into the Strasburg-Entzheim Airport in France in 1992. The crew inserted 3.3 into the flight management computer intending to fly a 3.3 degree glide slope. However, there was confusion about the current operating mode, and the pilots inadvertently inserted the command while in the vertical descent mode instead of the intended flight path control mode. This placed the aircraft in a 3300 feet-per-minute descent, causing it to crash in surrounding foothills. These pilots were highly trained with years of experience, and yet they failed to correctly interpret how to effectively communicate intent and the automation designers never anticipated that operators would make such a mistake.

Because of the existence of both low-level and high-level message sets in STANAG 4586 which allow the operator to both manually take over but also transition to higher levels of automation, the likelihood that mode confusion will be a serious problem is high. It is important to note that this is already a significant problem in all UAV operations but will likely be exacerbated by mode changes between low and high levels. Moreover mode confusion is not a state only operators find themselves in, it can also occur for designers of systems. Personnel who implement STANAG 4586 need to be sure they fully understand the modes, the phases of flight they are likely to occur in, and possible entry and exit strategies.

Information overload

The most significant human factors concern for operators of UAV systems is workload management. Research and actual operations have unequivocally demonstrated that operators under high workload do not achieve optimal performance, and commit more errors which can lead to the loss of entire platforms. If not designed with workload concerns in mind, STANAG 4586 HCIs could easily overwhelm controllers, which is generally true of all UAV interfaces. For example, the STANAG 4586 Implementation Guide contains a Command and Control Interface (CCI)-HCI matrix that maps data elements to specific displays. In this matrix, there are 812 instances of discussions of overlays that can go on displays as well as over 1400 data elements that should be displayed (which do not include vehicle and mission specific information that will need to be included). While all of the elements may be needed at some point, expecting an operator to remember where to find all the data elements, when, under what conditions, etc. will lead to unacceptable workload and potential mission and system failure.

More specifically, any UAV ground control system that requires operators to exert low-level control as depicted in the inner most loop in Figure 3 will significantly increase operator workload, potentially to unacceptable levels. In addition, low-level control like flying a vehicle requires significant training time. Furthermore, workload will increase any time operators must exert low-level control in any phase of flight, even in mission planning. While STANAG 4586 must necessarily address both low and high-level control to allow the participation of legacy systems, it is critical for human workload that the migration of system architectures be towards more autonomy, with less human intervention needed in the lower control loops.

Message/display integration

One concern in the implementation of interoperability standards is that designers will "stovepipe" data elements, which means they will take requirements literally and simply provide the functionality without consideration of the need for both data integration and the need to capitalize on human perceptual strengths through graphical presentation of data. Lack of information integration in displays was a contributing factor in the human operator mistakes in the Three Mile Island nuclear accident as well in the unintentional downing of an Iranian commercial airline by the *USS Vincennes*. In both cases, lack of consideration for the operators need to assimilate information across a variety of sources was not considered by designers, which was exacerbated in a time-critical environment, much like what UAV operators face.

Moreover, aural and visual warning integration is critical across sub-systems, especially in terms of priority handling for health and status warning messages. As seen in Figure 3, health and status monitoring is the activity that must occur constantly regardless of the control loop but it is not independent of the other phases. For example, real-time mission replanning should and could take into account various system capabilities (e.g., projected low fuel state, degraded cameras, etc.). However, if vehicle-specific information (such as operating parameters, limitations, checklists, etc.) cannot be integrated with mission planning information that comes from multiple, potentially external sources, then the workload of the operator will increase. Moreover, intelligent health and status monitoring is needed particularly for multiple emergencies and/or system failures that guides the operator through problem resolution. Thus the HCI and any related decision support will need to incorporate information that is both vehicle specific, and information from external C4I systems.

Feedback

Providing adequate feedback to operators is one of the most important HCI design principles and it is critical that changes of vehicle/payload/mission states are effectively communicated. Generally STANAG 4586 treats feedback as a binary condition, e.g., notify operator if an event has occurred or not. For example, there are a series of messages (700: handover failure, 800: mission upload failure, 900: upload aborted/rejected) that simply communicate to the operator that a condition has occurred or not. Unfortunately this information is of little help. What the operator needs is why the failure occurred and what the next steps are for resolution. Operators should not be forced to determine any of the recovery parameters and then recall detailed procedures, i.e., electronic checklists should be used to the greatest extent possible. Moreover, the use of trend displays for health and status monitoring is critical for allowing operators to build and maintain situation awareness but ensuring that both vehicle-specific message sets and operational message sets are integrated to accomplish this goal remains to be seen for interoperable ground control stations. While currently STANAG 4586 allows for vehicle specific generic messages and the correct display of vehicle specific information falls to the CUCS and VSM developers, display overhead, increased operator workload, and mode confusion are likely to result in complex off-nominal scenarios.

Future distributed architectures

The primary purpose of STANAG 4586 is to promote interoperability, and as previously discussed, it is important to consider the information requirements of the network of players. For interoperability to occur, the notion of "push" and "pull" for message sets defines how and who will have access to data across the network. As defined by STANAG 4586, "Push messages are broadcast either periodically or based on some event, but do not require a request to result in sending a message. Pull messages are messages that are generated in response to a request." Often individuals/agencies who are not in direct control (LOI 1/2/3) will need information on demand. For example, other operators/decision-makers not in the vehicle control loop may need payload configuration (300), stores management systems status (304), or communications relay status (305), all of which are currently defined as Pull messages. It isn't clear how this information would get to the correct individuals. While there is a process currently in place for formally changing or adding to STANAG 4586, (known as a document change proposal, DCP), in order to support situation awareness for all stakeholders in the network, it is important that all push/pull requirements are well understood and access provided.

CONCLUSION

As of now, STANAG 4586 is the only standard that exists to promote interoperability across a command and control network of unmanned vehicles. Such standards are needed to truly "flatten" the battlefield such that communications between allied forces occur without the significant uncertainty and technical difficulties that exist in joint operations today. What flexibility is lost when designing a system to a standard can be countered both with reduced costs and reduced degrees of freedom for uncertainty. This is particularly true when humans are in a supervisory control loop because introducing a standard clearly identifies boundaries. Disambiguating the design space both aids in the design process as well as the operation of such complex systems. Furthermore a standard such as STANAG 4586 will be critical in achieving the concept of "network-centric operations" or "network-centric warfare." More standards similar to STANAG 4586 will be needed to promote the interoperability intra and inter command and control networks of the future.

However, as outlined in this paper, there are numerous potential human supervisory control problems for human-systems integration when an interoperability network is introduced. As illustrated in Figure 3, information exchange levels between the air vehicle and the operator will cognitively saturate operators unless the system is afforded greater autonomy and the operator is permitted to work at higher levels of command abstraction. UAS HCIs, particularly those that will be part of an interoperable network, must present operators with abstracted capabilities, i.e., controlling using high level goals rather than primitive, low-level commands. To achieve this goal-based form of UAS control, more work is needed to ensure advanced automated intelligent decision support tools (such as automated path planners and automated target recognition) that require both vehicle specific information from a disparate set of vehicles in addition to mission planning information will be supported in interoperability standards such as STANAG 4586. Moreover, while STANAG 4586 does not and should not specify HCI design criteria, significantly more work is need to determine, document, and formalize good HCI practices.

By its nature, the development of such an interface specification or standard requires assumptions about the systems and their role. While government agencies have been instrumental in the development and championing of the Standard, they have done so and continue to refine STANAG 4586 in a constantly evolving operational landscape based on extant technology and not necessarily looking to what are the likely needs in the next 10-15 years. However, to achieve greater flexibility and cost effectiveness with unmanned aerial systems (UAS), greater system intelligence will be needed to cope with increased levels of uncertainty than what is currently envisioned in the Standard today. However, greater autonomy does not come without a price and human authority must remain a core value of any design if the autonomous system is to provide useable capability [4]. The design of an autonomous UAS depends not only on the addition of "smart" technologies but equally on the Human Computer Interface (HCI) and the nature, timeliness and relevance of the information presented to the operator together with the level of control afforded over the capability.

ACKNOWLEDGMENTS

This work was sponsored in part by the Air and Weapon Systems Department, Dstl Farnborough.

REFERENCES

[1] Chairman, "STANDARDISATION AGREEMENT (STANAG) 4586: Standard Interfaces of UAV Control System (UCS) for NATO UAV Interoperability," vol. March, N. S. A. (NSA), Ed., 2 ed: NATO, 2005.

[2] DoD, "Network Centric Warfare: Department of Defense Report to Congress," Office of the Secretary of Defense, Washington, D.C. July 21, 2001 2001.

[3] T. B. Sheridan and W. L. Verplank, "Human and Computer Control of Undersea Teleoperators," MIT, Cambridge, Man-Machine Systems Laboratory Report 1978.

[4] J. T. Platts, "Autonomy in Uninhabited Air Vehicles," The Aeronautical Journal, The Royal Aeronautical Society, February, 2006.