Modeling multiple human operators in the supervisory control of heterogeneous unmanned vehicles

Citation: Mekdeci, Brian, and M. L. Cummings. "Modeling multiple human operators in the supervisory control of heterogeneous unmanned vehicles." In Proceedings of the 9th Workshop on Performance Metrics for Intelligent Systems - PerMIS 09, 1. Association for Computing Machinery, 2009.

As Published: http://dx.doi.org/10.1145/1865909.1865911

Publisher: Association for Computing Machinery

Persistent URL: http://hdl.handle.net/1721.1/81781

Version: Author's final manuscript: final author's manuscript post peer review, without publisher's formatting or copy editing

Modeling Multiple Human Operators in the Supervisory Control of Heterogeneous Unmanned Vehicles

Brian Mekdeci
Massachusetts Institute of Technology
77 Massachusetts Ave, Rm 33-407
Cambridge, MA, 02139
1-617-452-3044
mekdeci@mit.edu

M.L. Cummings
Massachusetts Institute of Technology
77 Massachusetts Ave, 33-311
Cambridge, MA, 02139
1-617-252-1512
missyc@mit.edu

ABSTRACT
In the near future, large, complex, time-critical missions, such as disaster relief, will likely require multiple unmanned vehicle (UV) operators, each controlling multiple vehicles, to combine their efforts as a team. However, is the effort of the team equal to the sum of the operator's individual efforts? To help answer this question, a discrete event simulation model of a team of human operators, each performing supervisory control of multiple unmanned vehicles, was developed. The model consists of exogenous and internal inputs, operator servers, and a task allocation mechanism that disseminates events to the operators according to the team structure and state of the system. To generate the data necessary for model building and validation, an experimental test-bed was developed where teams of three operators controlled multiple UVs by using a simulated ground control station software interface. The team structure and inter-arrival time of exogenous events were both varied in a 2x2 full factorial design to gather data on the impact on system performance that occurs as a result of changing both exogenous and internal inputs. From the data that was gathered, the model was able to replicate the empirical results within a 95% confidence interval for all four treatments, however more empirical data is needed to build confidence in the model's predictive ability.

Categories and Subject Descriptors
I.6.3 [**Computing Methodologies**]: Simulation and Modeling – *applications.*

General Terms
Performance, Experimentation, Human Factors.

Keywords
Discrete event simulation, human factors, modeling, team performance, supervisory control, unmanned vehicles.

PerMIS'09, September 21-23, 2009, Gaithersburg, MD, USA.

1. INTRODUCTION
Unmanned vehicles (UVs) are currently in use for numerous military operations, but they are also being considered for many non-military applications as well, including mining, fighting forest fires, border patrol and supporting police [1]. Currently, several human operators are required to control many of today's UVs, but futuristic systems will invert the operator-to-UV ratio so that one operator can control multiple UVs [2]. To accomplish this goal, the level of automation will have to increase such that operators will give high-level, supervisory instructions to the UVs instead of manual control [3]. However, previous research has shown that even under supervisory control, there is a cognitive limit as to the number of UVs a single human operator can effectively manage [4, 5]. Large, complex, time-critical missions, such as disaster relief, will likely exceed that limit and will require multiple operators, each controlling multiple UVs, to combine their efforts. Since such systems do not currently exist, many questions arise, including: (1) How many operators are necessary to achieve a set of mission objectives? (2) How should the operators combine their efforts in the most effective way? (3) Will the group performance be more than, equal to, or less than the sum of the individual contributions?

2. RESEARCH OBJECTIVE
The goal of this research is to develop a quantitative model of a team of human operators, each performing supervisory control of multiple unmanned vehicles, in time-critical environments. This model would allow stakeholders, such as vehicle designers and battlefield commanders, to vary input parameters, such as vehicle speed and number of human operators, in order to determine their impact on system performance.

3. PREVIOUS RESEARCH

3.1 Queuing Model of Supervisory Control of Unmanned Vehicles
Supervisory control of unmanned vehicles involves an operator handling intermittent events via an automated system by giving high-level commands to UVs. As such, supervisory control of unmanned vehicles has been previously modeled as a queuing system where the vehicles requesting assistance are regarded as users and the human operators are regarded as servers [6]. For instance, in a simple surveillance scenario whose timeline is shown in Figure 1, an unidentified contact suddenly emerges at time *t*. This event, labeled A, requires that the operator perform a task, in this case, assign an UV to the contact location for further

investigation. Since this event is not directly controllable by the operator or vehicle, it is considered to be an exogenous event to the system. Ideally, the operator would notice this event and start "servicing" it immediately by performing the associated task. However, because of inherent inefficiencies of human attention, the operator will inadvertently introduce a delay between the arrival of this event and the moment he starts to service it (marked by event B in the timeline). This delay is due to a combination of the Wait Time due to loss of Situational Awareness (WTSA) and the Wait Time due to Interaction (WTI) [4]. WTSA occurs when the operator is not aware that the event requires his attention, whereas WTI occurs when the operator has noticed the event, but has not measurably started the associated task yet (perhaps due to deciding between the right course of action from a number of options). Since it is extremely difficult to separate WTSA from WTI, the measured time between when an event emerges and when the operator starts the associated task (assuming the operator is not busy and has the resources available to service the event) will be considered WTOD – wait time due to operator delay. Cummings and Mitchell [4] have shown that this delay can be quite significant particularly when operators are controlling multiple vehicles simultaneously and have degraded situational awareness.

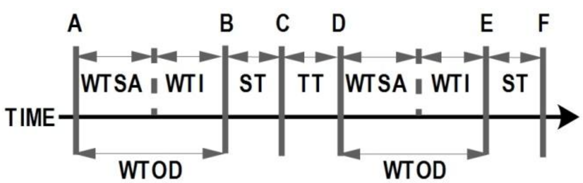

Figure 1: Timeline of events for simple UV scenario.

The task of assigning a vehicle to a location also takes a finite amount of time known as the Service Time (ST). At the moment when the operator finishes assigning a vehicle (C in Figure 1), that vehicle will begin to travel the assigned location. The time during which the vehicle is travelling is referred to as the Travel Time (TT) and in this scenario also represents the Neglect Time (NT) of the vehicle, since the vehicle acts autonomously during this period without requiring the operator's attention [7]. After some time, the vehicle will eventually arrive at the contact location, denoted by event D. Similar to the time between A and B, the vehicle must wait a finite period of time before the operator begins to interact with the vehicle's camera, denoted by event E. Finally, after another service time, the operator finishes identifying the contact (labeled event F) which may more may not spawn additional endogenous events, depending upon the scenario. If the final objective of the operator is to simply identify unknown contacts, then the difference in time between event F (when the final objective is met) and event A (when the contact emerged) is known as the Objective Completion Time (OCT). Since time is of the essence in many UV applications, the goal of many UV system designers and decision makers it to minimize the average OCT for a given scenario.

3.1.1 Multiple Event Handling

3.1.1.1 Wait Time due to Queuing

If an operator is busy interacting with a vehicle and another event emerges that requires the operator's attention, then that event must wait for the operator to become available. This additional time, not represented in Figure 1, is known as the wait time due to queuing (WTQ) since the event is considered to be in the queue for the operator's attention. Since vehicles tend to produce endogenous events (such as requiring new waypoints when they have reached the old ones), as the number of vehicles or exogenous events in the system increases, the probability of an event experiencing WTQ grows. Additionally, it has been shown that operators may take longer to respond to events as they emerge due to high workload and a loss of situational awareness [4]. Thus, as more events require the operator's attention, the OCT will continue to grow until it reaches an unacceptable level, at which point a team of multiple operators will likely be required.

3.1.1.2 Switching Strategy

If more than one event is in the operator's queue, the operator must select which event he will service next. There are several strategies an operator can use, including first-in-first-out (FIFO), highest-priority-first or even random selection. Switching strategy affects the total time tasks spend waiting for service not only because of the ordering of the tasks (queuing policy), but also because of the time required for the mental model change of the operator (switching cost) if the tasks are dissimilar [8]. It has been demonstrated that for operators of multiple, unmanned vehicles, the switching cost can be substantial [9].

3.2 Single Operator Discrete Event Simulation Model

Solving traditional queuing models can yield results of interest to the study of supervisory control such as the average time an event will spend waiting in a queue and server (operator) utilization. Although analytical solutions are possible for simple supervisory control systems, often the assumptions required for closed-form solutions, such as steady-state behavior and independent arrivals, are not met. Discrete event simulations (DES) overcome many of the limitations of analytical models by using computational methods that do not require such strict assumptions [10] and therefore allow a richer set of complex UV-operator systems to be modeled.

A single human operator controlling heterogeneous unmanned vehicles was successfully modeled using a Multi-UV Discrete Event Simulation (MUV-DES) model [8]. A Multi-UV, Multi-Operator Discrete Event Simulation (MUVMO-DES) model that builds upon this work, but also considers multiple operators combining their efforts, is the focus of this research. This new model consists of exogenous and internal inputs, operator servers and their interactions, and a task allocation mechanism that disseminates events to the operators according to the team structure and state of the system. The inputs to the model are both exogenous, such as the arrival rate of new contacts, and also internal, such as the length of time an operator spends interacting with a vehicle. These inputs are also stochastic due to the large amount of uncertainty in environmental conditions and human behavior.

4. METHODS

4.1 Multi-UV, Multi-Operator Discrete Event Simulation Model

Expanding the MUV-DES model to multiple operators required several new considerations, in particular a model of team communication, mutual performance monitoring and task allocation.

4.1.1 Modeling Communication

Geographically-disperse UV operators communicate through voice, chat or a combination of both. Voice communication is typically the fastest and allows operators the ability to control the UVs while simultaneously communicating via a headset. Voice communication is effective for small teams but can become problematic as the number of operators becomes large, due to multiple voice messages that occur simultaneously. Thus, voice communications are typically serial in nature, meaning only one operator can speak at a time. Chat messages allow operators to send messages to each other asynchronously and in parallel. Due to software's ability to parse text and apply sorting filters in real-time, chat communication often scales well with large teams. Chat messages also tend to be clearer than voice communication, in that they are not as susceptible to noisy communication channels, background noise, volume or operator accents. Furthermore, chat messages automatically create a real-time transcript of the communication, something that is typically not possible with voice. For the initial MUVMO-DES model, communications are assumed to be chat for data gathering purposes, but given the widespread use of chat by operational command and control personnel, this assumption also carries external validity. Modeling voice communications is left for future work.

4.1.2 Mutual Performance Monitoring

In addition to explicit communications, operators may also coordinate by mutual performance monitoring, recognized as one of the core components of teamwork [11]. Through a user interface, operators can typically view each other's vehicles and commands to gain situation awareness of what the team is doing. For instance, instead of explicitly communicating, an operator may take a quick look at the interface to see if any other operator's vehicles are already heading to a new contact before assigning their own. However, because this form of coordination is unilateral, teammates must make assumptions about the actions and intentions of other teammates which may or may not be valid.

4.1.3 Modeling Coordination

Communication and mutual performance monitoring can be represented by discrete endogenous events that the operators generate. For instance, in Figure 2, instead of servicing an event once it arrives (event A), an operator may choose to send a chat message to other operators by first starting a chat message, composing it for a finite period of time (labeled COORD) and then sending it before starting to service the task (event C). Similarly, an operator may perform a mutual performance monitoring task that also takes a finite period of time. However, if an operator is composing a chat message or monitoring the performance of other operators, then the operator is considered to be busy and as such, any event that is waiting for the operator's attention while he is communicating or monitoring will incur a WTQ for that period of time. This additional WTQ represents a quantitative measurement of the coordination cost (process loss) associated with the team performance.

The timeline shown in Figure 2 is a simple example of coordination but more complex coordination scenarios exist as well. For simple tasks, a single communication message may be all that is needed, such as claiming responsibility for a target that emerges. For more complex tasks, the communication may involve a conversation that spawns several iterations of communication messages. This initial model will only assume single communication messages and as such, will only be able to model simple coordination between the team members.

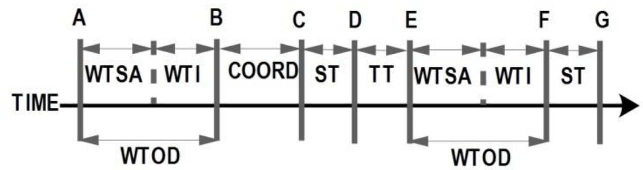

Figure 2: Timeline of events with coordination.

4.1.3.1 Coordination Strategies

Similar to switching strategies, an operator will also have a coordination strategy that dictates the type and timing of the coordination he will perform when faced with a task that can be serviced by more than one operator. One such strategy is to not coordinate at all, but this would require the team to have predefined roles and responsibilities (such as mechanistic teams) or run a high risk of task allocation errors. A task allocation error occurs when more than one operator or no operator attempts to service a particular task.

If an operator choose to coordinate her actions, she typically must choose the type of coordination first, i.e. whether or not to communicate, monitor or both. In addition to the type of coordination, the timing of the coordination is very important as well. A common strategy would be to coordinate first and then service the task. This type of coordination strategy is the least likely to incur task allocation errors. This coordination strategy was assumed for the initial MUVMO-DES model. However, other coordination strategies exist. For instance, an operator could service the task first and then send a courtesy message to other operators. This strategy allows the operator to give the fastest response to an event, but raises the possibility that another operator will also begin servicing the task before the first operator gets a chance to send the coordination message.

4.1.3.2 Team Structure and Task Allocation

Although the model was designed to be general and handle a variety of team structures, *mechanistic* and *organic* teams structures were chosen to be modeled initially since they represent two polar opposites of the organizational spectrum [12]. A mechanistic team is one where the operators have rigidly defined roles and responsibilities. For instance, when all of the vehicles of one type are assigned to one and only one operator, then that operator is given the full responsibility for performing the tasks that only that vehicle can do. If one of each vehicle type is allocated to each operator instead, then that team structure would be considered organic since any operator can perform any task that arises, provided that he has an appropriate vehicle available. Both team structures suffer from inefficiencies, or what Steiner [13] refers to as a "process loss" which is the differential between the performance of a team and the theoretical maximum achieved if the efforts of the individuals were combined ideally. In mechanistic teams, process loss occurs when task loads are uneven and some operators are too busy while others are idle. In organic teams, process loss occur when operators have to spend time coordinating how they will share the common queue and/or allocate the tasks amongst themselves in a sub-optimal manner.

Due to the clear task allocation roles, extending the MUV-DES model for mechanistic teams involved having a separate queue and server for each operator. Since each task was unique to an operator, every event that arose was automatically assigned to the appropriate operator.

For the organic team, a different task allocation mechanism was needed. Since the model is merely an abstraction of the actual scenario, the first attempt at an organic model randomly assigned the tasks to the operators based on who was available at that moment to service the event. If more than one operator was available, the event was randomly assigned to one of the available operators. If no operator was available, the event waited in a common queue (incurring a WTQ cost) until an operator became available. This form of modeling assumes that there will be no task allocation errors, i.e. one and only one operator will service or attempt to service any particular task. In real organic teams, this will likely only happen if the teams coordinate their actions through communication or mutual performance monitoring.

4.2 Data Gathering

The MUVMO-DES model utilizes stochastic processes to account for the uncertainty within the system. Therefore, random values are drawn for WTOD, service time, communication time, monitoring time, travel times and travel time in the model. These probability density functions (pdfs) need to be generated by binning empirical data into histograms and fitting an appropriate curve.

To generate the stochastic inputs necessary for model building and to validate the model's outputs against actual team performance metrics, real data must be gathered. Since there are no extant systems of teams of operators each controlling multiple unmanned vehicles, there is no "real world" data to collect. Hence, an experimental test-bed where teams of operators controlled multiple UVs was specifically developed and experimental trials were conducted to gather the data used for model building and validation.

Figure 3: Main display of the ground control interface.

4.2.1 Experimental Test-Bed

The experimental test-bed consisted of a video game-like simulation of unmanned vehicle control by a team of operators. The simulation included three ground control stations, with one subject assigned to each station.

4.2.2 Ground Control Interface

Subjects interacted with the ground control stations via a computer monitor display using standard keyboard and mouse inputs. The main display of the ground control station featured three sections – a large map, a chat panel and a system panel (Figure 3). The map represented the geographical area that the operators were responsible for, as well as all the vehicles under their control and contacts that they needed to handle. Contacts and vehicles were represented using MIL-STD-2525B icons [14] and the operators assigned vehicles to contacts by clicking on the map interface with the mouse. The operators were also able to communicate with each other via instant messaging within the chat interface window. Operators would type messages into the chat, which would then appear on all the other operator's chat panels instantly. Chat messages were labeled with the operators unique IDs, which corresponded to the labels for each operator's vehicle icons. In addition to the map and chat display, there was also a system panel where the system would occasionally send messages to a particular operator, such as a confirmation message that the operator had assigned a particular vehicle to travel to a particular location.

4.2.3 Tasks

Each mission scenario required a team of operators to "handle" contacts that appear intermittently over the map. To do this, the team of operators needed to perform both assignment and payload tasks.

4.2.3.1 Assignment Tasks

Assignment tasks required the operators to send their vehicles to the contacts on the map as they emerged. Once assigned, the UV would start to travel to that particular contact location on the map in a straight line and would continue until either the vehicle reached its assigned destination or the operator re-assigned the vehicle elsewhere. There were no obstacles on any of the maps and no path-planning required.

Although assignments were done by individual operators, they can be considered a "team task" since the operators had to coordinate their assignments to ensure that one and only one vehicle was assigned to each and every contact. Furthermore, subjects were instructed that vehicles should be chosen in the interest of minimizing travel times, i.e. typically the closest available vehicle to the contact location.

4.2.3.2 Payload Task

Once a vehicle reached a contact, the operator performed a simple task by interacting with the vehicle's payload. This task was unique to the vehicle and contact type, but involved either visual identification (e.g., where is the red truck in the parking lot?) or a simple hand-eye coordination task. Since all three vehicles were aerial of some sort, all payload tasks involved a birds-eye view of the terrain. An example of a hand-eye coordination task is shown in Figure 4 where the operator must destroy a contact by centering the crosshairs over a stationary target on the ground and pressing the fire button three times. The difficulty in this task was that the crosshairs are subject to jitter due to the motion of the UV. The other hand-eye coordination task involved dropping aid packages to victims on the ground. This task was similar to the destruction task except that the crosshairs were steady but the projectiles were slow-falling and susceptible to the wind. Thus, players had to compensate for a light north-east wind, for instance, by aiming packages slightly to the southwest of the target location and pressing the drop button once. Payload tasks are considered an "individual task" as they do not require any coordination or assistance from any of the other operators.

Figure 4: Missile firing payload task.

4.2.3.3 Scenario Objectives

The objective of each scenario was to identify all unidentified contacts and either rescue them (if friendly) or destroy them (if hostile) as quickly as possible. There were three vehicle types, one that handles each type of contact (unidentified, friendly, hostile) exclusively. Although any UV of the appropriate type could be assigned to a contact, only the first vehicle to start the payload task could successfully complete it. When a contact first appeared on the map, it was always of the unidentified type, which required a scouting UV (Type A). Once the scouting UV arrived, the operator performed a visual identification task which transformed the contact from unidentified to either hostile or friendly. If the contact was identified as being hostile, a tactical UV (Type B) was sent by an operator to the contact location to destroy it via the missile firing task. Similarly, if an unidentified contact was identified as being friendly, a rescue UV (Type C) was sent by an operator instead to drop aid packages to the contacts' location, thereby "rescuing" the contact. The time a contact spent in the system, from the moment it arrived, until the moment it was successfully handled, was the objective completion time. Since a scenario consisted of multiple contacts, the Average Objective Completion Time (AOCT) was the metric of interest, where the average was simply the mean of all the OCTs for that scenario.

4.2.3.3.1 Design of Experiments

A 2x2 repeated measures experiment was conducted where the independent variables were team structure (mechanistic, organic) and the inter-arrival time of unidentified contacts (constant, erratic). Ten teams of three participants each completed all four treatments. The order of trials was counter-balanced and randomly assigned to the teams. An alpha value of 0.05 was used for significance.

4.2.4 Independent Variables

4.2.4.1 Inter-Arrival Times of Exogenous Events

Previous research has demonstrated that optimal UV operator performance occurs when the operator has a utilization lower than 70% [15]. Thus, all scenarios were designed to have an operator utilization of about 50%, meaning that operators spent approximately 50% of their time, on average, performing assignment or payload tasks. This was achieved in pilot studies by fixing the payload tasks and manipulating the number of exogenous events and their inter-arrival times until the average operator utilization was about 50%.

The experimental trials had a total of 16 exogenous events (unidentified contacts emerging). The time between successive exogenous events (the inter-arrival time) was 30 seconds for the constant treatment. For the erratic factor level, the inter-arrival times were generated from a bimodal distribution where the means of the modes were set at 75 seconds and 225 seconds from the start of the trial, with a standard deviation of 15 seconds. In both the constant and bimodal treatments, the first exogenous event always appeared at time 0, thus only 15 events were drawn from the bimodal distribution for the erratic condition. The inter-arrival of exogenous events was varied between constant and erratic to determine if team structure had an effect on how operators performed under different task load distributions.

4.2.5 Participants

Participants were recruited via e-mail and paper advertisements and through word-of-mouth. All of the participants were between the ages of 18 and 35, with the mean age being 21.7. Some participants had military, video game or previous UV experiment experience. Due to scheduling concerns, some teams were composed of individuals who knew each other while most teams were composed of individuals who were randomly assigned. The level of inter-personal relationships between team members (stranger, casual acquaintance, friend, romantic, etc) was not recorded.

4.2.5.1 Training

Prior to the experimental trials, the participants completed an individual 20-minute PowerPoint® training session. Afterwards, the participants completed two practice scenarios (one mechanistic and one organic) as teams, each one taking about 10 minutes to complete. Thus, the total training time was approximately 40 minutes.

5. RESULTS

The order of the trials was checked to determine if a learning factor occurred across the four team sessions. Given that the training time was minimal, and previous research has shown that four or more training sessions is needed for teams to achieve stable performance [16], testing order was of concern, and showed a significant effect ($F(3, 24) = 4.12$, $p=.02$). Most teams did worse on the first trial, regardless of the treatment, than on subsequent trials (Figure 5). Thus, the final statistical model included a two factor, repeated measures ANOVA with blocking on the trial order.

Team structure was significant ($F(1, 24) = 1.484$, $p < 0.01$), with mechanistic teams performing better than organic teams overall, although there was no significant difference when the inter-arrival rate was erratic. Mechanistic teams performed worse when the inter-arrival rate was erratic as opposed to constant ($t(15.8) = 2.47$, $p = 0.03$). However the inter-arrival rate had no significant effect on the organic teams. The inter-arrival rate by itself was not significant, but the interaction of the independent variables was ($F(1, 24) = 10.47$, $p = 0.04$).

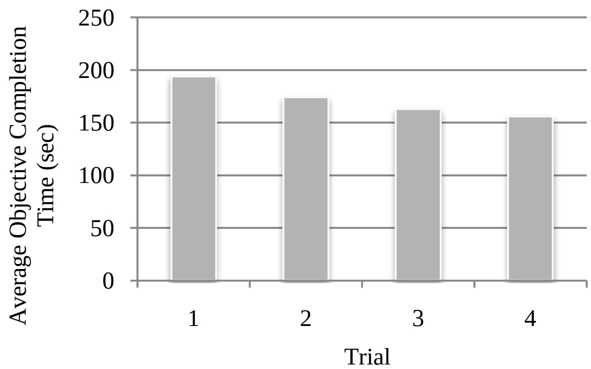

Figure 5: Effect of AOCT vs trial order.

5.1 Model Results

The model was run 1000 times for each treatment condition. For the organic team, the model predictions were within the 95% confidence interval of the empirical results for all four treatments (Figure 6). Since the mechanistic teams did not have to coordinate their actions due to their rigid role structure, they were initially modeled without any communication or monitoring behavior. In the erratic inter-arrival condition, the model predictions for the mechanistic team was within the 95% confidence interval, however for the constant inter-arrival condition, the model's predictions were low (Figure 6).

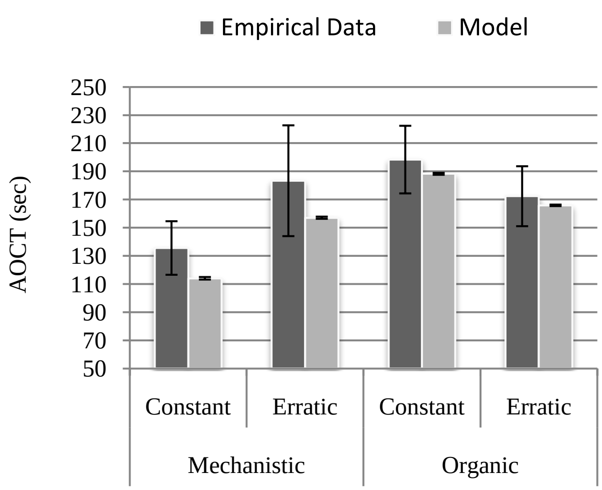

Figure 6: Initial empirical results.

Upon further investigation of the experimental transcripts, the mechanistic team did communicate and monitor each other's actions, even though it was not necessary. Thus, a coordination strategy similar to that used by the organic team was implemented in the mechanistic model and new outputs were generated. Not surprisingly, the additional cost associated with coordination increased the OCT of the mechanistic team. Thus, with the coordination strategy implemented in both teams, the model predictions were within the 95% confidence interval for all four treatment conditions (Figure 7).

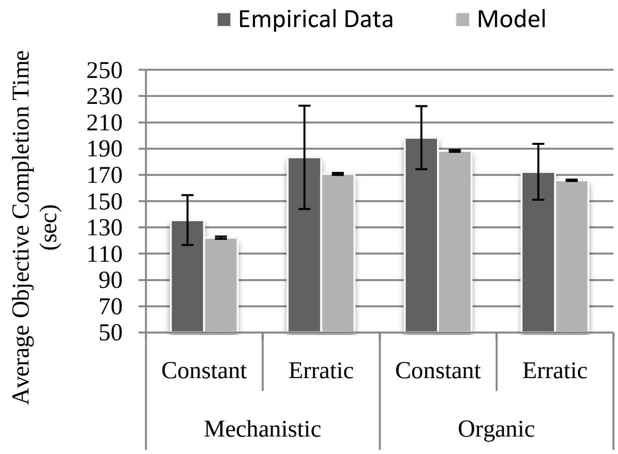

Figure 7: Revised empirical results.

6. DISCUSSION

It was not surprising that the mechanistic teams performed worse under erratic inter-arrival times than they did when the inter-arrival times were constant, since the erratic inter-arrival times caused events to arrive in batches, thereby increasing the queues. However, it was interesting that there was no significant difference in the performance of the organic team under the different inter-arrival rates of exogenous events. This suggests that even though events arrived in clusters during the erratic inter-arrival treatment, the organic team was able to handle the workload spike without increasing the AOCT. This suggests that the organic team is more robust to environmental uncertainty than mechanistic teams due to their flexible structure and the ability to spread tasks across the team.

It was predicted that mechanistic teams would perform better than organic teams, which they did, but not necessarily for the same reasons. Originally, mechanistic teams were thought to have an advantage over organic teams because they did not incur coordination costs. As shown in the results, mechanistic teams do incur coordination costs and without taking these costs into consideration, the performance predictions are too low in the constant inter-arrival case. This is interesting because the communications are theoretically unnecessary. However, this highlights the importance of understanding the intrinsic need for communication between team members, even if it is not necessary. Future work should look at how to mitigate such communication overhead.

So, if mechanistic teams are also incurring coordination costs, how are they still managing to perform better overall than organic teams? The answer to this question perhaps lies in the fact that the empirical data used to generated the pdfs for the different sources (e.g. travel times, WTOD, service times) was separated into the four different treatment conditions. Although there was no statistically significant difference between the values and the differences could be attributed to sampling error, there were small differences in nearly every input condition. Since the OCT is the sum of all of these individual times, then these differences (or errors) combine into a statistically significant result.

Other factors may play a role as well, such as the switching strategy of the operators. The switching strategy assumed for all of the operators was FIFO, although in many cases, operators did not adhere to this strategy. Thus, future analysis should determine the actual switching strategies observed in the experimental trials and implement those instead.

Another issue is that statistical significance for data such as WTOD was difficult to obtain due to a number of factors. First, the sample size of the experiment was small (n = 10) but this is not unusual for team studies since it takes multiple participants to form a single experimental unit. Increasing the sample size should reduce the standard error of the experimental results. Additionally, previous research has shown that UAV teams do not reach asymptotic performance levels until after they have completed around four sessions together [16]. Although this is likely to be highly contingent upon a number of factors such as the difficulty of the task, the inter-operability required for success and the length of the sessions, it does seem to be consistent with our results. Thus, to further reduce variability in the experimental results, additional practice sessions should be added. Finally, the experiment was not controlled for the skill level or the relationships of the individuals. Factors such as age, video game experience and military background could have had an effect on individual performance. If a reduction in the variability of the team's performance is desired, then future experiments could select for and block on particular individual traits. However, teams of futuristic UV operators may be just as diverse as the sample population, particularly if they are composed of individuals from different agencies or even nations operating via an interoperability standards [17]. These operators may have different levels of training, skills and attitudes which may result in significantly different levels of individual performance. Thus, it is not necessarily a flaw in the experimental design to have diversity in regards to the individual traits, as it can be argued that such diversity will be likely in future UV systems.

7. FUTURE WORK

The model in this paper has successfully replicated the results of experimental trials, but it has not been used to predict the performance of teams in hypothetical situations. Future work will look at developing the model to predict the performance of teams in new scenarios and then verify those results empirically. One such scenario could be if the teams had an additional member or decision support tool that aided in task allocation. While the mechanistic teams performed better than organic teams overall, the fact that the mechanistic teams were more sensitive to variations in the environment suggests that this team architecture may not be ideal for volatile environments such as those found in command and control settings. If an organic team had the benefit of a leader or decision support tool, then its coordination costs might drop significantly, whereas a leader or decision support tool would likely have little or no effect on a mechanistic team. Thus, the team model could be updated to see just how much of a performance difference one could expect by having a leader or decision support tool in both team structures.

8. ACKNOWLEDGMENTS

This research is sponsored by the Office for Naval Research and the Air Force Office of Scientific Research. We would also like to acknowledge M.I.T. undergraduates Tony MacDonald, Kyle Miller, Anrae Jamon Motes, Ben Sedat, Jieyun Fu and An V. Ho for their programming contributions to this project and post-doctorate associate Birsen Donmez for her assistance with statistical analysis.

9. REFERENCES

[1] A. Ollero, & Maza, I., *Multiple Heterogeneous Unmanned Aerial Vehicles*, Berlin: Springer, 2007.

[2] J. Franke, V. Zaychik, T. Spura *et al.*, "Inverting the Operator/Vehicle Ratio: Approaches to Next Generation UAV Command and Control."

[3] M. L. Cummings, S. Bruni, S. Mercier *et al.*, "Automation architecture for single operator, multiple UAV command and control," *The International Command and Control Journal*, vol. 1(2), 2007.

[4] M. L. Cummings, and P. J. Mitchell, "Predicting controller capacity in remote supervision of multiple unmanned vehicles," *IEEE Systems, Man, and Cybernetics, Part A Systems and Humans*, vol. 38, no. 2, 2008.

[5] S. R. Dixon, C. D. Wickens, and D. Chang, "Mission control of multiple unmanned aerial vehicles: A workload analysis," *Human factors*, vol. 47, no. 3, pp. 479, 2005.

[6] M. Cummings, C. Nehme, J. Crandall *et al.*, "Predicting operator capacity for supervisory control of multiple UAVs," *Innovations in Intelligent UAVs: Theory and Applications*, vol. 70, 2007.

[7] D. Olsen, and M. Goodrich, "Metrics for evaluating human-robot interactions."

[8] C. Nehme, "Modeling Human Supervisory Control in Heterogeneous Unmanned Vehicle Systems," Aeronautics & Astronautics, Massachusetts Institute of Technology, Cambridge, MA, 2009.

[9] M. Goodrich, M. Quigley, and K. Cosenzo, "Task switching and multi-robot teams."

[10] J. Banks, and J. Carson, "Discrete-event system simulation," Prentice-Hall, 1984.

[11] E. Salas, D. E. Sims, and C. S. Burke, "Is there a "Big Five" in Teamwork?," *Small Group Research*, vol. 36, no. 5, pp. 555-599, October 1, 2005, 2005.

[12] T. Burns, and G. Stalker, *The management of innovation*: Oxford University Press, USA, 1994.

[13] I. D. Steiner, *Group process and productivity*: Academic Press New York, 1972.

[14] "Common Warfighting Symbology," *MIL STD 2525B*, Defense, ed., 2005.

[15] B. Donmez, C. Nehme, M. Cummings *et al.*, "Modeling situational awareness in a multiple unmanned vehicle supervisory control discrete event simulation," *Journal of Cognitive Engineering and Decision Making, Special Issue on Computational Models of Macrocognition (in review)*, 2008.

[16] N. J. Cooke, P. A. Kiekel, and E. E. Helm, "Measuring team knowledge during skill acquisition of a complex task," *International Journal of Cognitive Ergonomics*, vol. 5, no. 3, pp. 297-315, 2001.

[17] M. Cummings, A. Kirschbaum, A. Sulmistras *et al.*, "STANAG 4586 Human Supervisory Control Implications," *Air and Weapon Systems Dept, Dstl Farnborough & the Office of Naval Research*, 2006.