Operator Choice Modeling for Collaborative UAV Visual Search Tasks

Luca F. Bertuccelli, *Member, IEEE*, and Mary L. Cummings, *Senior Member, IEEE*

***Abstract*—Unmanned aerial vehicles (UAVs) provide unprecedented access to imagery of possible ground targets of interest in real time. The availability of this imagery is expected to increase with envisaged future missions of one operator controlling multiple UAVs. This research investigates decision models that can be used to develop assistive decision support for UAV operators involved in these complex search missions. Previous human-in-the-loop experiments have shown that operator detection probabilities may decay with increased search time. Providing the operators with the ability to requeue difficult images with the option of relooking at targets later was hypothesized to help operators improve their search accuracy. However, it was not well understood how mission performance could be impacted by operators performing requeues with multiple UAVs. This work extends a queuing model of the human operator by developing a retrial queue model (ReQM) that mathematically describes the use of relooks. We use ReQM to generate performance predictions through discrete event simulation. We validate these predictions through a human-in-the-loop experiment that evaluates the impact of requeuing on a simulated multiple-UAV mission. Our results suggest that, while requeuing can improve detection accuracy and decrease mean search times, operators may need additional decision support to use relooks effectively.**

***Index Terms*—Decision theory, man machine systems, unmanned aerial vehicles.**

I. INTRODUCTION

AN IMPORTANT aspect of ongoing and envisaged unmanned aerial vehicle (UAV) missions is the visual search task, in which operators are responsible for finding a target in an image or a video feed. Due, in part, to advances in networked sensors, military analysts are becoming increasingly overwhelmed with the volume of incoming UAV imagery (both full motion video and static images) [1]. Given the future Department of Defense vision of one operator supervising multiple UAVs, the amount of incoming imagery to be analyzed in real time will grow [2]. Moreover, with recently announced wide-area airborne sensors such as Gorgon Stare and Argus which can generate up to 64 images per single UAV camera concurrently, there is an urgent need to develop efficient approaches for human analysis of UAV-generated imagery [1], [3].

Given the complex interactions between the human and the automated sensors in these UAV missions, models of the human operator are necessary in order to develop more appropriate decision support systems (DSSs) that account for operator decision-making inefficiencies, such as increased wait times for vehicle selection and loss of situation awareness [4]. Mathematical models for human operators interacting with multiple UAVs have been developed using a queuing framework [5], [6], where external tasks are generated from an underlying stochastic process and the human supervisor, modeled as a server, services the stream of tasks. While analysis of realistic multi-UAV missions is analytically intractable, discrete event simulation (DES) of operator queuing models has been used to generate accurate performance prediction of experimental results [6]. Operator models have also been developed for human information aggregation using two-alternative-choice (2-AC) models [7]–[11] and visual search formulations [12]–[22].

In difficult search environments, operators searching imagery in multi-UAV environments may desire more choices than determining if a target is present or absent in an image. More specifically, operators may seek additional information in order to find the target, possibly through another visit later on in the mission, or they may choose to ignore a task because there is not sufficient information to make a confident assessment. We hypothesized that, instead of a two-choice model, operators would be better served by having a third option of reevaluating a search task at a later time by requeuing the image and taking another glance via a relook. Throughout this paper, we make a distinction between the choice of requeuing, which is the abandonment of the current search task, and a relook, which is an additional glance at a previously searched image.

While it is straightforward to implement a requeue option in a multi-UAV simulator, the effect of providing requeue and relook capabilities must be investigated experimentally since there is a potential for undesirable effects such as increased operator workload. Furthermore, these capabilities could be operationally important in minimizing collateral damage and reducing errors and have been studied in the context of optimal stopping [23] and inspection problems [24]. While these works showed promising results in assessing the informational value of an additional look, these studies are limited in two main ways. First, previous work related to the speed–accuracy

Manuscript received December 28, 2010; revised June 28, 2011 and December 28, 2011; accepted January 18, 2012. Date of publication May 1, 2012; date of current version August 15, 2012. This work was supported in part by the Michigan/Air Force Research Laboratory Collaborative Center in Control Science and in part by ONR Grant N00014-07-1-0230. This paper was recommended by Associate Editor D. Kaber.

L. F. Bertuccelli was with the Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, Cambridge, MA 02139 USA. He is now with the United Technologies Research Center, East Hartford, CT 06108 USA (e-mail: luca.bertuccelli@utrc.utc.com).

M. L. Cummings is with the Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, Cambridge, MA 02139 USA (e-mail: missyc@mit.edu).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSMCA.2012.2189875

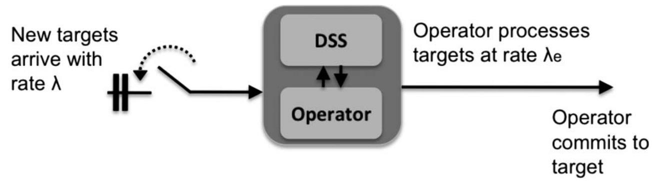

Fig. 1. Queuing model for the human operator in which new targets arrive at a known rate λ and they are processed at a rate λ_e.

tradeoff shows that operator accuracy may degrade with time, due to either vigilance effects attributed to fatigue [25] or task difficulty [26]. Furthermore, complex missions such as those involving multivideo visual search tasks or missions that require both planning and searching tasks can increase operator workload [6].

This effort makes three novel contributions in presenting a choice model for a search task with requeues. First, we develop a retrial queue model (ReQM) for visual search tasks that includes the possibility of requeuing difficult images and pose ReQM as a variation of a retrial queue with feedback [27]–[30]. We next develop a DES with ReQM (DES-ReQM) that embeds operator models derived from previous experimental data of a simulated multi-UAV mission. We then present results in the predicted performance of multi-UAV visual search tasks using DES-ReQM. We build on previous work [31] by discussing the results of a human-in-the-loop experiment that confirm the predictions made by DES-ReQM, as well as presenting experimental observations of operator behavior in requeuing and relooking tasks. We conclude with a discussion on the implications for requeues and relooks with an actual operator in the loop.

II. Operator Modeling

Queuing models for human operators have been previously proposed in the context of air traffic control, where the human operator is treated as a serial controller, capable of handling one complex task at a time [5]. These queuing models of operators have been recently extended to account for operator workload and attention inefficiencies in the context of human supervisory control of multiple UAVs [6]. Fig. 1 shows a general queuing model for a multi-UAV supervisory control problem for the visual search task. Search tasks are generated by a Poisson process at an average rate λ, and the human operator (with possible help from a DSS) services the tasks at a rate λ_e. In complex tasks, operators may dedicate themselves only to a single task at a time, allowing the incoming tasks to accumulate in the queue. The visual search task initiates when the operator begins examining the image feed once the UAV reaches the target, and concludes with a decision on the target location.

A. Decision Models

An important feature of an operator queuing model is that a submodel is needed to understand how humans accumulate information and ultimately make detection decisions in search tasks (the "Operator" block in Fig. 1). One common formulation uses a 2-AC framework [7]–[11], [32]. 2-AC models originate from hypothesis testing models [7] and characterize information accumulation as a stochastic diffusion process. It can be shown [33] that sufficient statistics of the diffusion model can be summarized by two random variables that can be measured empirically: the probability of choosing one of the alternatives P and the mean response time $\bar{T}$. For the visual search task in this paper, P is the probability of detection.

Previous work in the visual search literature has also attempted to provide some insight in human response times and accuracy in a speed-accuracy tradeoff setting. For example, signal detection theory (SDT) has been used to mathematically characterize human performance in visual search tasks. In most experiments, subjects are shown a sequence of images, and both detection time and accuracy are measured. Earlier work has proposed generating a receiver operating characteristic for humans to understand the relationship between correct detection and false alarms [20], while other work extends the SDT framework to include attention [34]. King *et al.* [35] used realistic imagery and assumed an SDT model to describe the performance of the subjects. Waldman *et al.* [36] developed an empirical model for visual detection with search, based on identifying parameters of a probabilistic model, such as search time and accuracy. Sperling and Melchner [37] investigated visual search with attention more specifically looking at reaction times and investigating a so-called attention operating characteristic. More recently, Huang and Pashler [38] have considered attention in visual search.

Extensive work in visual search has also emphasized the use of probabilistic models that relate mean decision time to the mean time to search [13]–[15], [18]–[20]. Recent work has also moved beyond the mean decision times and analyzed the role of parameter identification for parameterization of the search time distributions [39], [40]. While the characterization of the search is also important from a cognitive science perspective, the work in this paper does not address the low-level details of how the search is accomplished but rather seeks to understand and quantify the effect of sequential searches within the context of supervisory control.

B. Motivation for Requeues

We used experimental data obtained from a multi-UAV simulator including visual search tasks to determine the relationship between detection probability and search time in search tasks performed by single operators in multi-UAV simulated missions [6]. The search tasks contained imagery obtained from Google Maps, and the participants were instructed to maximize the number of targets found over the course of the mission. In multiple target searches, this previous work observed that subject probability of detection degraded with time since difficult searches required additional cognitive effort from the operators [41] and operators had to switch between planning and searching [6]. We used these data to determine that the empirical mission probabilities of detection decreased with increased search time and hypothesized that the increased likelihood of mistakes arises because people are forced to make

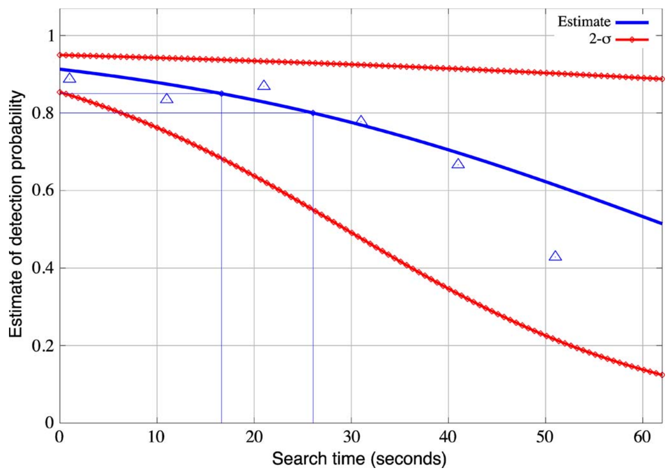

Fig. 2. Detection probability decreases as a function of time due to shared attention between planning and visual search tasks: The best estimate is the solid line, the 2-σ regions are shown as diamond-etched lines, and the maximum-likelihood estimate is shown as triangles. The vertical lines show illustrative examples of search times for detection probabilities of $P_d = 0.85$ and $P_d = 0.80$.

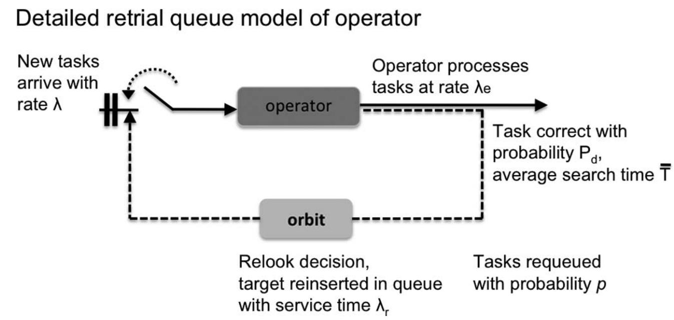

Fig. 3. In contrast to the queuing model in Fig. 1, retrial queuing allows for operator requeues in difficult tasks, which are then placed in an orbit queue. Targets are then reinserted in the available task list at some later time, and the operator can choose to search or abandon them.

a choice on a search task in order to move on to waiting search tasks.

Fig. 2 shows the probability of detection modeled using a logistic regression for a visual search task obtained from previous experimental data in multi-UAV simulated missions [6] and is consistent with previous vigilance literature [26], [41]. The logistic model will be explained further in Section III, but given the decrease in operator accuracy with time, it appears that there could be a benefit to requeuing the current search task (and possibly abandoning it in the absence of new sources of information). While one of the key benefits of a requeue is that it frees the operator to pursue other searches, particularly since the queuing model assumes that tasks are continually arriving, an additional benefit of a requeue is that a search task could be investigated at some later time in the mission via a "relook." Note that, when search tasks are requeued, they are simply reinserted in the queue, and there is no explicit provision for how a requeued task is searched again (e.g., first-come first-served (FCFS) policy). For example, an operator may choose to take another look at a requeued task after having explored other tasks or never search the task again for the remainder of the mission.

III. Retrial Queuing Model ReQM

In order to account for requeuing in a queuing framework, this section first describes the formulation for requeuing a task using retrial queues [27], [28] and describes the ReQM developed for this effort. Fig. 3 shows a visualization of ReQM. Just like in conventional queues for human operator models, ReQM treats the human as a server [6], and if the operator is free to initiate an available visual search task, the task is shown to the operator and can be serviced immediately. If the operator does not wish to complete the initiated task and wishes to delay it to some other time (which could lead to never seeing it again), the task is inserted in a so-called *orbit* pool. ReQM is a slight variation of the retrial queues in [29] and [30] but differs fundamentally by attempting to model how requeues and target detections are made by the operator (which are not addressed in [29] and [30]). Additional details of ReQM are provided in the following main components: 1) task arrival and service rates; and 2) operator choice model.

A. Arrival and Service Rates

ReQM assumes that new tasks follow a Poisson arrival with a rate λ and that new tasks are serviced by the operator at a rate λ_e. Note that, for queuing models of visual search tasks performed with UAVs, the UAVs are allocated in the environment to maximize the total number of targets found. Therefore, λ_e has a strong dependence on numerous mission-specific factors, but the principal drivers are operator search time and, in the case of a multi-UAV setting, vehicle routing policy that likely requires operator intervention [42]. Under certain arrival and service rates, queue instability can occur, in which the number of outstanding targets will grow unbounded over time. Vehicle routing policies have been developed to provide guarantees under which queue instability can be prevented [42], but it is unclear whether these guarantees will hold with actual operators in the loop, particularly if these operators have the capability of requeuing targets.

B. Choice Model

The choice model is the underlying mechanism under which the operator can make a detection decision (e.g., whether there is a target in the image or not) or decide that a task needs to be requeued.

1) Detection Decision: We abstract the operator choice model into a detection probability and search time distributions. For the detection probability, we derived a logistic regression model from previous experimental data [6]. The operator is assumed to make correct detections with probability

$$P_d(t_s) = \frac{1}{1 + \exp(\hat{\beta}^T \mathbf{t})} \quad (1)$$

where $\mathbf{t} = [1, t_s]$, t_s denotes the search time, and $\hat{\beta}$ is a vector of parameters obtained from the experimental data. In distinction to the work in [43], the detection probability is a nonstationary quantity and is negatively dependent on the search time.

Search time distributions can likewise be estimated from previous experimental data regarding the visual search task in a simulated multi-UAV experiment [6]. We found that the lognormal distribution in (2) is a good approximation for the search time distribution, where $\bar{T}$ and σ^2 are the mean and variance of the search times

$$f(t_s;\bar{T},\sigma^2)\propto\exp\left(\frac{-\left(\log(t_s)-\log(\bar{T})\right)^2}{2\sigma^2}\right),\qquad t_s>0. \quad (2)$$

2) Requeuing Decision: If the operator is not willing to make a detection decision, then the operator can choose to requeue the task. However, the requeuing policy describing how the operator decides to requeue tasks may depend on a number of factors, including the total amount of time spent searching for a target, the total number of remaining tasks, and target arrival rates. As a first approximation, ReQM assumes that an operator will requeue the target with some probability p, which is the probability that the search time exceeds some critical search time T_{rl} (additional details on how T_{rl} is chosen are provided in Section IV).

In summary, the operator choice model in ReQM assumes that the operators are going to either make a correct detection, make an incorrect detection, or ask to requeue the task. Search times are distributed according to $f(t_s;\bar{T},\sigma^2)$, and in the event of a detection decision given a realization t_s from this search time distribution, the operator makes a correct detection with probability $P_d(t_s)$.

C. ReQM Analysis and the Need for Simulation

ReQM is an initial attempt to represent how a repeated visual search task with a task requeuing option can be properly formalized using retrial queues. However, analysis of this model is difficult without human-in-the-loop experimental data, as it is unclear how frequently subjects decide to requeue tasks, and previous work in retrial queuing theory does not provide insight into these choices for human operators.

In addition, even if we understood how operators requeue targets, it would be difficult to analyze ReQM in closed form since real models for retrial queues may deviate from some of the common assumptions necessary for analytical tractability. In ReQM, for example, requeuing invalidates the assumption of independent arrivals. Furthermore, Choi and Park [29] assume that a task in the orbit queue can only be serviced if the nominal queue is empty. This is not a suitable representation for the multiple-UAV relook problem, since a target can be requeued regardless of the remaining outstanding visual search tasks. Queuing theory is also concerned with queue stability, in the sense that the number of tasks does not grow unbounded over time, which may not be a valid assumption when human performance is considered.

IV. DES of ReQM: DES-ReQM

While it will be the topic of future work to investigate whether the analytical methods from retrial queuing theory may be applicable, a method for admitting less restrictive assumptions can be addressed by using DES. First, where analytical methods are not available for analyzing a queue in closed form, DES can help provide insight of the queue transient properties. Second, DES can be used for tuning the appropriate set of parameters to be used for human-in-the-loop experiments, such as determining appropriate task arrival rates. The ultimate goal of the DES environment in this effort is to provide a high-fidelity simulation of the experiment, and this section discusses a DES model of ReQM, DES-ReQM, which is composed of three main parts: an environmental module, a routing policy module, and a requeue policy module.

A. Environmental Module in DES-ReQM

The environment is assumed to be a bounded region, populated with stationary targets that are generated according to a Poisson process with arrival rate λ. Without loss of generality, for the kinds of single-operator multi-UAV missions envisioned for this work, we assume that the low-level vehicle control loops are closed by an onboard autopilot and that low-level planning problems (such as satisfying turn rate constraints on UAVs) are not the responsibility of the operator but that of appropriate low-level control algorithms.

B. Operator Planning Module in DES-ReQM

In modeling the operator planning policy, we make the assumption in DES-ReQM that operators allocate UAVs to targets according to a policy that routes UAVs to the targets that are nearest geographically. While the current research is investigating the role of different routing strategies [42], we will assume this greedy approach.

We are interested in a mission objective that maximizes the total number of targets found (N_F) out of the total number of possible targets in the environment (N_T)

$$J_F = N_F/N_T. \quad (3)$$

N_F is a function of numerous operator-specific parameters (such as target difficulty, search time, and requeue policy), but in the multi-UAV setting, it also has a strong dependence on the routing policy for the vehicles. For example, operators could increase vehicle travel time by assigning a vehicle to visit a distant target, rather than allocating vehicles to service nearer tasks.

Upon reaching the targets, the UAVs are assumed to loiter around the target and initiate a visual search task only when the operator has chosen an available UAV. For the ReQM, UAVs initiate a visual search task according to an FCFS policy (in which the first UAV to reach the target initiates the search task first), which is a common assumption made in vehicle routing problems [42]. The search times were modeled by using the search time distributions from previous experiments performed in the Research Environment for Supervisory Control of Heterogeneous Unmanned Vehicles (RESCHU) multi-UAV simulation environment [6]. Realizations of the search times for the ith task t_s^i are generated by sampling a new random

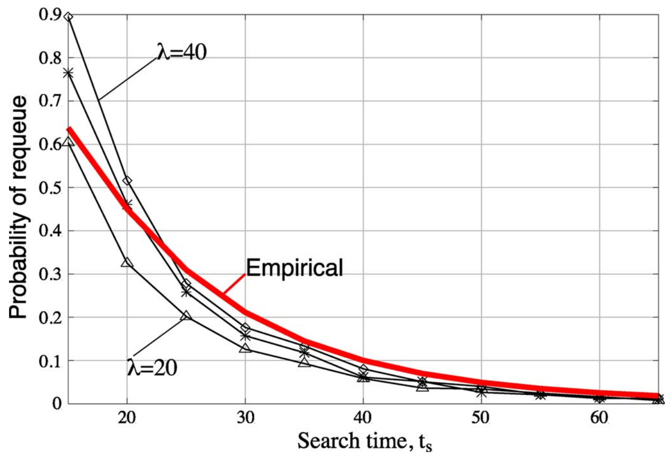

Fig. 4. Trend of $(\triangle, *, \diamond)$ empirical relook probabilities agrees with that of (solid red line) theoretical relook probability for three different arrival rates chosen for testing in DES-ReQM: $\lambda = \{20, 30, 40\}$ (in seconds per target).

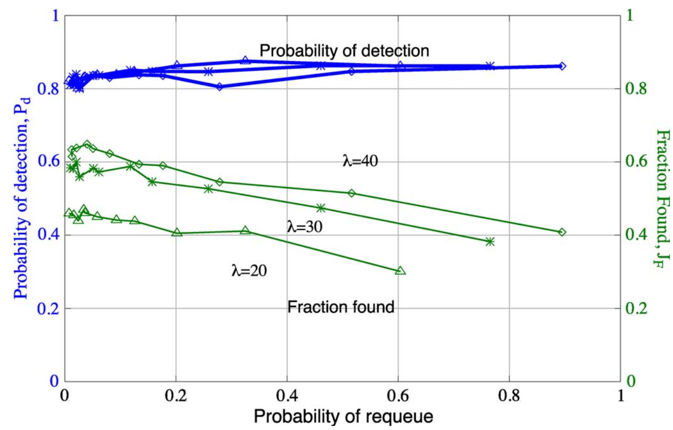

Fig. 5. Detection probability and fraction found J_F versus requeue probability for three different arrival rates. Lower arrival rate decreases J_F from 0.62 to 0.42 as relook probability increases to 0.8, but detection probability increases from 0.82 to 0.86.

number from the lognormal distribution in (2) with mean $\bar{T}$ and variance σ^2

$$t_s^i \sim f(t_s; \bar{T}, \sigma^2). \tag{4}$$

In turn, the search outcome is generated by the realization of the random variable $P_d(t_s^i)$ from the logistic regression of (1). For the human-in-the-loop experiment discussed in the next section, we determined, from previous experimental data [6], that the logistic regression parameters are given by $\hat{\beta} = [-2.300, 0.037]$. For the search time distributions, we found that $\log(\bar{T}) = 3.1$ and $\sigma^2 = 0.6$.

C. Requeue Module in DES-ReQM

The requeuing model in DES-ReQM assumes that a task is requeued by the operator if the realization of the search time t_s^i exceeds a critical time T_{rl}. If the realization of the search time is less than T_{rl}, the task is not requeued, and a probability of correct detection is calculated using (1). For example, suppose that the critical time is chosen as $T_{rl} = 25$ s, and a new search time is realized from the distribution in (4). Suppose that this realization was $t_s^i = 27$ s. In this case, the task is automatically requeued by the system when $T_{rl} = 25$ s has elapsed. When a requeue takes place, the task is inserted in the orbit pool, and a new route is calculated for any available UAV in the simulation. If, however, the realization of the search task was $t_s^i = 23$ s, the task would not be requeued, and the probability of a correct detection would be calculated with (1).

The critical time T_{rl} was varied as a simulation parameter and discretized in the interval $T_{rl} \in \{20, 25, 30, \ldots, 60\}$. This interval was chosen since it described over 95% of the support of the search time data from previous experiments. Note that the theoretical requeue probability $\bar{p}$ for each of the critical times T_{rl} can be found with the following integral (which is the red line labeled "Empirical" in Fig. 4):

$$\bar{p}(T_{rl}) = \Pr(t \geq T_{rl}) = 1 - \int_0^{T_{rl}} f(t_s|\mu, \sigma^2)\, dt_s. \tag{5}$$

Unfortunately, the integral is not available in closed form, but numerical routines can evaluate the cumulative distribution function of the lognormal distribution.

D. Simulation Results

This section presents simulation results of the performance using the previously developed operator choice models analyzing one hundred 10-min-long simulated UAV missions. In this setting, we analyzed the detection probability (P_d) and the fraction found [J_F, given by (3)].

Targets were modeled given with three distinct average target arrival rates $\lambda \in \{20, 30, 40\}$ (in seconds per target) that, given previous human-in-the-loop experimental data, were arrival rates representing different task loads. Fig. 4 shows the agreement between the theoretical predictions relating the requeue probability with the search time T_{rl}. For example, choosing a relook search time of $T_{rl} = 25$ s implies that the probability of requeue will be on the order of $p = 0.3$. Recall from (5) that, as the search time T_{rl} threshold increases, we intuitively expect the probability of requeuing to decrease since operators will have more time to make decisions on the presence or absence of the target.

By varying the time T_{rl}, it is possible to, in turn, investigate the effect of requeuing on mission performance. Fig. 5 shows the impact of varying the probability of requeuing on the detection probability and fraction found J_F, averaged over the 100 Monte Carlo simulations. Fig. 5 shows that DES-ReQM predicts that tasks that were requeued with probability $p = 0.3$ for arrival rates of $\lambda = 30$ s/target resulted in a fraction found of $J_F = 0.5$, while requeuing with probability $p = 0.78$ resulted in a fraction found of $J_F = 0.4$ Predictably, the increase in requeue probability decreases the fraction found J_F since operators do not have enough time to find new targets. On the other hand, the probability of detection under all three arrival rates increases from 0.81 to 0.86, thereby demonstrating the potential for an improvement in the overall probability of detection by virtue of moving on to less challenging tasks.

(a)

(b)

Fig. 6. RESCHU (a) interface and (b) search panel. (a) RESCHU interface: UAVs (shown as a blue bell shape) are directed to fly to (red diamonds) locations with targets while avoiding (yellow circles) risky areas. (b) RESCHU search panel during a search task showing the (top right) timer counting down and the relook button.

In summary, DES-ReQM shows that there is an important tradeoff between maximizing the fraction of targets found and ensuring a high overall accuracy in target detection. Therefore, requeues may be beneficial in the context of improving probability of detection, but a human-in-the-loop experiment is needed to investigate operator requeuing strategies and the effect on mission performance. We discuss the relation between experimental and simulation results in the later sections.

V. Relook Experiment

Using the results from the DES-ReQM simulations, an experiment was conducted with the objective of investigating the performance of a retrial queue when users had an available requeue option. The experiment was performed in RESCHU [6], a simulation specifically tailored to investigate human-in-the-loop interaction with multiple UAVs. A typical RESCHU interface is shown in Fig. 6(a). A single operator is tasked with handling N UAVs in an environment where targets are nonmoving but appear at random intervals. When a UAV (shown as a blue bell shape) reaches a location of interest (shown as a red diamond), a visual search task is initiated by the operator in the top left panel of the interface, with a magnified version of the search panel shown in Fig. 6(b). Note that, in the visual search task panel, the operator can zoom in and out the display while panning the image. In addition, the operator can query the system with the "Query" button to find out how many residual search tasks still need to be processed in the mission.

A. Experimental Objective

In this experiment, a single operator was responsible for the coordinated search of an area using six homogeneous UAVs. The objective of this experiment was to maximize the fraction found, which was explained to the participants as the total number of targets found out of the total number of possible targets in the environment (3). This experiment had two treatments: 1) a relook mode (a "within-subject" treatment) and 2) a timer condition to induce artificial time pressure (a "between-subject" treatment).

In the first treatment, the operators were tested under three different relook modes.

1) No relook (NR): Operators did not have the ability to relook. Operators were required to commit to the location of a target before returning to the UAV planning task.
2) Relook with consent (RWC): Operators had the option of requeuing at any time, but after T_{rl} seconds, a flashing message was displayed on the search screen to suggest to the operator to requeue.
3) Relook without consent (RWOC): Operators had the option of initiating a requeue at any time, but after T_{rl} seconds, the target was automatically requeued.

The second treatment involved the use of the timer and was inserted to provide operator feedback on how much time had been spent searching. Previous work has shown that time pressure can cause different operator strategies, so the two experimental conditions were with and without a timer [4].

B. Search Task and Visualization of Requeued Targets

The search images were obtained from Google Maps, and aerial views of different scenes were presented to the operator. Each scene contained a target that needed to be found by the operator, and hence, errors were only in the class of missed detections. Different images were presented to the operator. Fig. 7 shows two examples of search tasks, where Fig. 7(a) shows a search task with the objective of finding the fighter jet (located in the lower right corner of the image). Fig. 7(b) shows an example of a search task requiring the detection of a helicopter landing pad in a cluttered environment. The tasks were randomly placed in the mission, and pilot tests were used to populate the image database to ensure sufficient diversity.

In the RWC and RWOC conditions, a target became available for assignment T_r seconds after it was requeued. Tasks that were requeued changed color for $T_r = 15$ s to show that they were not available for search. When they became available

(a)

(b)

Fig. 7. Examples of imagery used in the interface. The search task in (a) requires that the participant finds the fighter in the image, while the task in (b) requires a search for the helicopter pad.

again, the tasks were changed to their initial color but were enclosed by an orange circle to inform the operator that they had been previously visited. The relook counter was incremented by one if the targets with an encircling orange circle were searched again.

When operators were ready to make an assessment of the target, they right clicked the image on their best estimate of the target location, and a small menu appeared. One menu item was "Submit," which, when selected, was evaluated as correct or incorrect by the software. If the participant mistakenly clicked the screen, a "Cancel" menu allowed the participant to return to the search.

C. Experimental Participants and Procedures

A total of 36 participants took part in the experiment (8 female and 28 male): Thirty participants were 18–25 years old, four participants were 25–35 years old, and two participants were older than 35 years. After a 10-min training session, the participants were randomly assigned a relook mode (NR, RWC, or RWOC) and a timer setting (timer or no timer) and participated in three sessions with a fixed timer setting and a counterbalanced order of relook modalities. The participants performed the first 20-min test session, took a rest break, repeated a second 20-min session (with another relook setting), took another rest break, and concluded with a third 20-min experiment session. The order of the trials was randomized and counterbalanced on the timer conditions between subjects. The experimental parameters were chosen as follows: $N = 6$ UAVs, $\lambda = 30$ s/target, $T_{rl} = 25$ s, and $T_r = 15$ s.

VI. Experimental Results

A mixed model analysis of variance of 2 (timer or no timer) $\times$ 3 (NR, RWC, or RWOC) repeated measures was used for statistical analysis. Three participants had to be excluded since they were outside the 3-σ interval for the fraction found given by (3).

A. Performance With Relook Modalities

The results in this section present the probability of detection, the mean search time, and the fraction found for the timer condition and decision support modalities. Recall that, since there is a target present in every image, then the probability of a missed detection is the only error for these search missions.

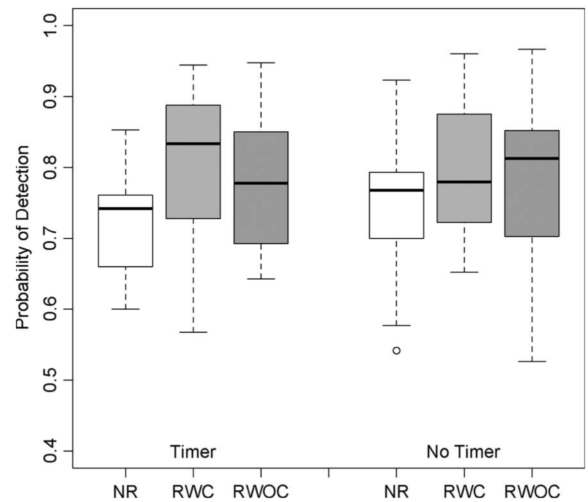

Fig. 8. Probability of detection increases most with a timer using RWC and most with RWOC.

Fig. 8 shows that the probability of detection improved, on average, from 0.72 (with NR) to 0.80 (with RWC), for the case when a timer was used (left). Furthermore, when no timer was used (right), the probability of detection increased, on average, from 0.74 (with NR) to 0.78 (with RWOC). For the overall probability of detection, the mode was significant at the $\alpha = 0.05$ level, with RWOC and RWC showing better detection than NR, with $F(2, 62) = 6.674$ and $p = 0.002$. *Post hoc* comparisons using a Tukey honest significant difference (HSD) test indicated that RWC and RWOC did not differ from each other, and there was neither a significant main effect of timer nor a significant interaction between timer and modality. In particular, the change in the mean of the detection probability in the NR mode (mean: 0.73; standard deviation (SD): 0.09) was significantly different $(p = 0.001)$ from that in the RWC mode (mean: 0.80; SD: 0.10). The change in the mean of the detection probability in the NR mode (mean: 0.73; SD: 0.09) was also significantly different $(p = 0.03)$ from that in the RWOC mode (mean: 0.78; SD: 0.11). These results demonstrate that providing the operator with requeuing choices improved their accuracy in a statistically significant manner. Practically, an improvement of accuracy from 0.72 to 0.80 is significant for UAV operations, since it decreases the likelihood of collateral errors.

Next, we investigated the effect of operator search times for the tasks. The operator search times for repeated looks in the RWC and RWOC modes were aggregated together to ensure a fair comparison to the NR mode and are shown in Fig. 9 and hence are representative of the total time spent searching the images. It can be seen that the mean search time decreased from 21.0 s (with NR) to 19.1 s (with RWC) for the case when a timer was used (left). Furthermore, when no timer was used (right), the mean search time decreased from 22.5 s (with NR) to 19.4 s (with RWOC). For the overall mean search time $(\alpha = 0.05)$, the mode is significant: $F(2, 62) = 7.032$; $p = 0.002$. The timer

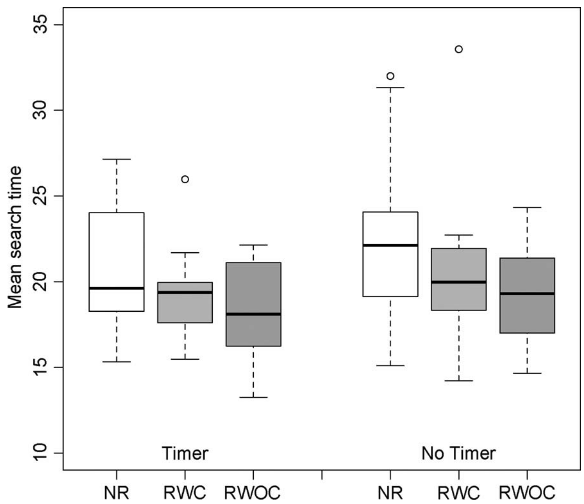

Fig. 9. Mean search time decreases for both RWC and RWOC.

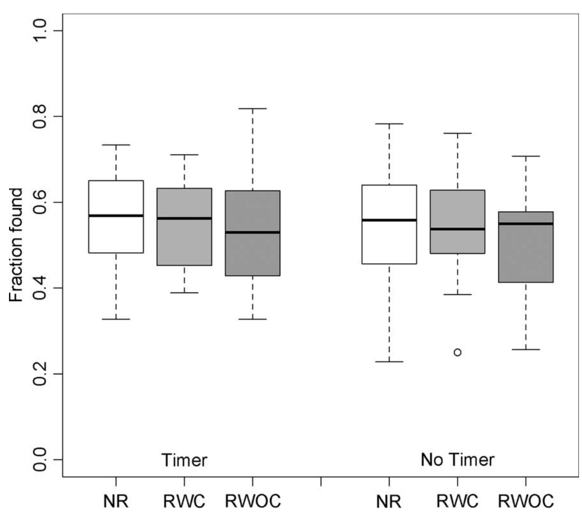

Fig. 10. Downward trend in the fraction found is not statistically significant.

effect and the interaction between mode and timer were not significant.

Post hoc comparisons using a Tukey HSD test indicated that the change in the mean search time in the NR mode (mean: 21.82; SD: 4.38) was significantly different ($p = 0.02$) from that in the RWC mode (mean: 19.79; SD: 3.49). The change in the mean search time in the NR mode (mean: 21.82; SD: 4.38) was also significantly different ($p < 0.001$) from that in the RWOC mode (mean: 18.91; SD: 2.87). These results may also have practical multi-UAV significance, where improvements in the speed of detection of the location of an adversary can have dramatic consequence for mission success.

Lastly, the objective of this experiment was to maximize the correct fraction that is the quotient of the total number of targets found and the total number of targets possible. The results for the correct fraction for each requeue mode and timer condition are shown in Fig. 10. As anticipated by the DES in Section III, there appeared to be a decrease in the fraction found as relooks are employed: The mean fraction found decreased from 0.55 (with NR) to 0.51 (with RWOC). This change in the correct

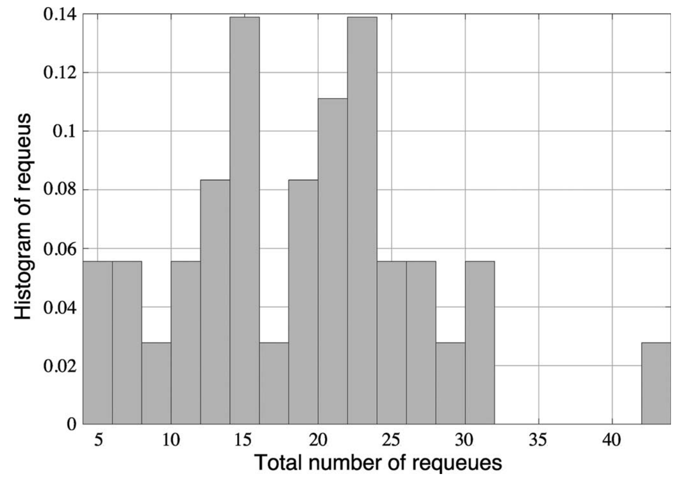

Fig. 11. Total number of requeues made by the 36 participants in the RWOC and RWC modes. Note the outlier with a total of 44 requeues.

fraction was not statistically significant with respect to the mode [$F(2, 62) = 1.254$; $p = 0.29$] or timer [$F(2, 62) = 0.03$; $p = 0.974$]. However, a Pearson correlation revealed that, as the probability of a relook increased, the fractions of targets found decreased ($r = -0.71$; $p < 0.001$). This highlights the cost of such an action in that the more frequently operators elected to reinsert tasks in the queue, the less time they had to prosecute other targets.

B. Behavioral Analysis of Use of Relooks and Requeues

The next step in the analysis was quantifying the number of times that requeuing was actually implemented by the participants (Fig. 11). A total of 697 requeues was made by the subjects. The 12 subjects that had the highest number of requeues accounted for 47.9% of the total requeues made by all subjects. Interestingly, the subject that requeued the most targets had previous actual UAV experience and may have been more inclined to relook at the targets for operational considerations. In the RWC mode, a total of 269 requeues was made, with 149 requeues occurring after the time limit $T_{rl} = 25$ s had expired. This means that 55.3% of the participants ignored the recommendation made by the decision support algorithm and chose to continue searching the image. In turn, 44.7% of the participants anticipated the automation's prompt. In the RWOC mode, a total of 428 requeues was made, with 274 requeues being implemented by the decision support algorithm because the participants searched longer than $T_{rl} = 25$ s. Out of all the participants in the RWOC mode, 10 out of 36 people never requeued voluntarily (the algorithm requeued the tasks for them all the time). Of the people that did requeue voluntarily at least once, they did so, on average, 41.6% of the time.

Target relooks occurred less frequently than requeues, and the summary statistics are shown in Table I. Interestingly, there was a statistically significant difference between the probabilities of requesting a relook between the RWC and RWOC modes [$F(1, 35) = 13.46$; $p < 0.001$]. Out of all the visual search tasks, a total of $N = 3499$ was analyzed in the first look, $N = 311$ in the second look, $N = 72$ in the third look, and

TABLE I
SUMMARY STATISTICS OF RELOOKS

Number of relooks	0	1	2	3
Number of tasks inspected	3499	311	72	17
Probability of error	22.7%	41.5%	63.2%	85.7%
Number of requeues	524	128	34	11

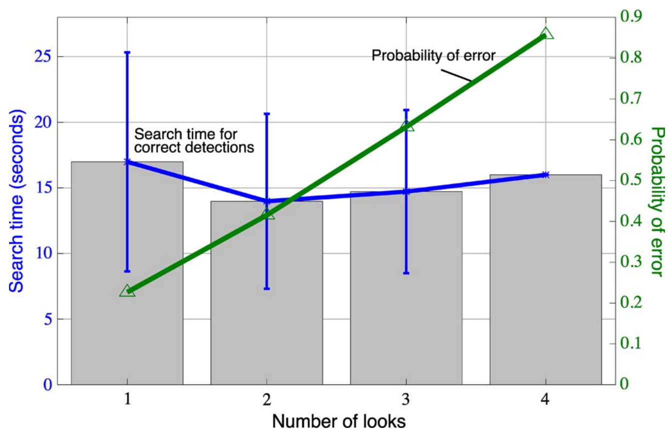

Fig. 12. Mean search time for targets that were detected decreases with additional relooks. However, the probability of error increases, suggesting that the subjects were not receiving additional information with each relook.

$N = 17$ in the fourth look. Note that the first look takes into account not only tasks that were searched successfully but also tasks that were skipped (either with or without consent). In the first searches, a total of $N = 2978$ target detections was made (either correct or incorrect), with a total of $N = 524$ requeues made after having investigated the targets. Out of the total looks that were made, only 17.8% of the tasks were looked at more than once. The discrepancy between the number of requeues and the number of looks is attributed to the fact that some targets were skipped and never relooked again. This implies that operators were using the requeuing function predominantly to move on from difficult images.

To investigate further the existence of a benefit to relooking at targets later in the mission, we show the mean search times for correct detections and probability of errors associated with the different numbers of looks in Fig. 12. Note that the mean search time for correct detection slightly decreased from 17 to 15 s with an increased number of relooks. This trend was also visible for targets that were missed. However, the overall probability of error increased as a function of the number of looks: from 22.7% in the first look to 41.5% in the second look, 63.2% in the third look, and only 85.7% in the last look.

A logistic regression model was generated by using the number of relooks as a categorical variable, and the overall effect of the number of relooks was statistically significant with a χ^2 test ($\chi^2 = 64.5$; $df = 4$; $p < 0.001$). The difference in the coefficients of the logistic regression model from one look to two looks was statistically significant ($\chi^2 = 5.7$; $df = 1$; $p = 0.02$), but the difference from two to three looks and from three to four looks was not significant. Nonetheless, this apparent increase in error, coupled with the fact that only 17.8% of the targets were looked again two or more times, hints further at the possibility that the operators that were being presented with the same imagery later in the mission were pressured to make an assessment and frequently made this assessment erroneously. Additional discussion on this observed effect is presented in Section VII.

Increased error with relooks of the same image has profound ramifications for supervisory control of UAV missions, because it suggests that operators may be willing to make a mistake to avoid repeating the same searches. Furthermore, it also suggests that the perceived benefit of the requeuing methodology is the freedom to keep exploring other tasks, rather than being forced to make a choice on a difficult image. (Recall that targets that were requeued by the operators were enclosed by an orange circle so that they could be clearly observed by the operator.)

C. Subjective Assessment of Confidence and Workload

Subjective assessment of the different requeue models is important, since the operator must ultimately accept or reject the recommendations set forth by such automated DSSs. In developing DSSs for complex mission planning involving the visual search task, two key subjective assessments are confidence and workload.

The subjective confidence assessment was a reflection of how accurately people felt about their performance in detecting the targets in the image. Upon completion of a search task, participants were asked about their subjective confidence on the accuracy of their detection. A five-point Likert scale was used, where 1 indicated "not very confident" and 5 indicated "very confident." The probability of detection was averaged for each condition that the participant performed, and a correlation between confidence and average probability detection yielded a Pearson correlation coefficient of $r = 0.42$ $(p < 0.001)$, demonstrating that participant confidence was significantly correlated with their performance.

The three different modalities showed a statistically significant difference in confidence, at the $\alpha = 0.05$ level $[F(2, 62) = 4.39;\ p = 0.02]$. However, neither the interaction between mode and timer was significant. *Post hoc* comparisons using a Tukey HSD test indicated that the change in the mean of the self-assessed confidence in the NR mode (mean: 3.79; SD: 0.59) was significantly different $(p = 0.01)$ from that in the RWC mode (mean: 3.95; SD: 0.48), indicating that operators felt that they performed better in the RWC mode.

Workload was measured in two distinct ways. In the first subjective method, participants were asked to rate their own level of workload. For the self-assessed workload and at the end of each search task, participants could enter 1 (not very loaded) to 5 (very loaded) on a five-point Likert scale. Changes in the self-assessed workload across the mode were not statistically significant at the $\alpha = 0.05$ level $[F(2, 62) = 1.92;\ p = 0.15]$. In addition, the interaction between mode and timer was not significant $[F(2, 62) = 1.10;\ p = 0.34]$, and neither was the timer treatment significant $[F(1, 31) = 2.10;\ p = 0.16]$. In the second method, interval utilization (or percent busy time) was used to gauge the objective workload in the experiment [44], [45]. Interval utilization has been shown in previous work

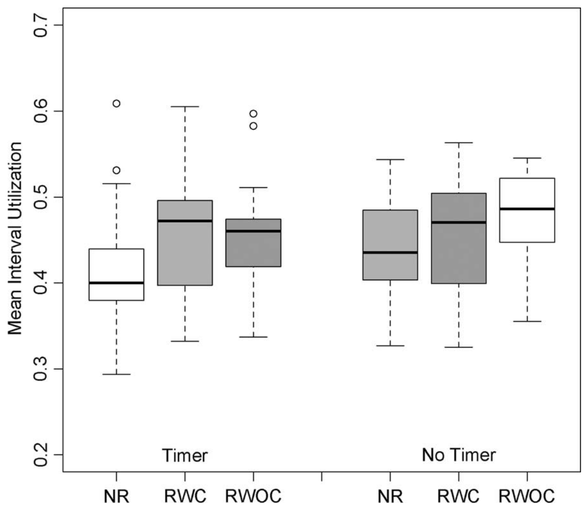

Fig. 13. Interval utilization increases with RWC and RWOC.

to be a reliable assessment for workload [6]. For the mean interval utilization (Fig. 13), participant interaction time with the user interface was normalized by the time between detection tasks and averaged for each participant. The increase in mean utilization across the mode was statistically significant at the $\alpha = 0.05$ level [$F(2, 62) = 6.49$; $p = 0.003$]. The interaction between mode and timer was not significant [$F(2, 62) = 0.57$; $p = 0.57$], and the timer treatment was also not significant [$F(1, 31) = 0.55$; $p = 0.47$]. While the differences in participant interaction time were statistically significant, the small practical difference in percent utilization (from 43% in the NR mode to 46% in the RWC mode) may have made this difference imperceptible to the subjects and hence may be a possible reason for the lack of statistical significance in the subjective assessment.

Post hoc comparisons using a Tukey HSD test indicated that the change in the mean of the interval utilization in the NR mode (mean: 0.43; SD: 0.07) was significantly different $(p = 0.02)$ from that in the RWC mode (mean: 0.46; SD: 0.07). Also, the changes in the mean interval utilization in the NR mode (mean: 0.43; SD: 0.07) and RWOC mode (mean: 0.46; SD: 0.07) were significant $(p = 0.002)$. While interval utilization increase is expected in the RWC and RWOC modes since the operators had an additional tool to use, utilizations on the order of 40%–70% are still within the acceptable range for human supervisory control [46].

D. Comparison to DES Model

Finally, we validate the predictions made by DES-ReQM by comparing the results obtained in the DES and the experiment. For these simulation comparisons, the NR case was simulated by assuming that T_{rl} was a very large number such that the empirical probability of a requeue was zero in all the simulations, while RWOC was the case described earlier for the case of $T_{rl} = 25$ s. The RWC case in Table I was determined by finding the corresponding probability of requeuing that was observed in the experiment. Table II presents the results comparing the use of requeues in terms of the fraction found metric as predicted by DES-ReQM and obtained experimentally. The first column shows the mode condition (NR, RWC, and RWOC), the second column shows the mean and SD of the fraction found predicted by DES-ReQM, the third column shows the experimental results, and the fourth column shows the power analysis. Note that, in order to compare the RWC experimental and simulated conditions, we had to first determine the empirical average probability of requeuing which we found to be $p = 0.38$. Nonparametric Mann–Whitney U-tests showed no statistical difference between the predictions made by the DES-ReQM and the experiment for all the three conditions: $z = 1.41$ and $p = 0.16$ for NR, $z = -0.51$ and $p = 0.11$ for RWC, and $z = 1.21$ and $p = 0.23$ for RWOC.

TABLE II
FRACTION FOUND DES-ReQM COMPARISON TO EXPERIMENT

	DES-ReQM		Exp		p-value	Power
	μ	σ	μ	σ		
NR	0.58	0.14	0.55	0.14	p=0.16	0.27
RWC	0.53	0.16	0.53	0.12	p=0.61	0.05
RWOC	0.55	0.15	0.51	0.13	p=0.23	0.42

VII. Discussion on the Use of Requeues

In this experiment, the probability of detection and the mean search time improved with the presence of the requeue option, whether mandated or not. The cost of such requeues meant somewhat increased objective workload (but no increase in subjective workload), and the more the participants that accessed the relook feature, the fewer the targets that they were likely to find. Such results highlight the cost-benefit issues surrounding any new decision support tool in that it can often provide benefit, but there are also possible negative consequences if such a tool is invoked too often.

One of the interesting results from this experiment was that detection probability increases as subjects are provided with the capability to requeue. However, relooks actually increase the probability of making an error. This seemingly counterintuitive result arises from the total number of targets that were searched in the NR mode compared to the RWC and RWOC conditions. In fact, subjects in the NR mode made an average of 32.6 total searches, while 31.2 and 30.8 were made in the RWC and RWOC modes. However, in the RWC mode, only 25.8 and 22.0 total decisions were made, and of these decisions, the respective error probabilities were 19.2% and 19.1% for RWC and RWOC. In contrast, the NR mode had an error rate of 27.8%, suggesting that the benefit of the relook was to allow people to skip difficult targets. Nonetheless, it appears that people felt implicit pressure to make a decision for repeated targets. Thus, this research suggests that, when there is not new information in an additional glance, allowing operators to requeue (i.e., skip) a target, but not relook at it, may be a more effective strategy when tasks are arriving stochastically. This result has broader implications for the value of information of an additional look, as well as teaming of operators, where it may be advantageous to requeue skipped targets for other teammates to avoid the increased likelihood of a mistake.

VIII. Conclusion and Future Work

This paper has developed a choice model for an operator performing visual search tasks generated from multiple unmanned vehicles. Using previous experimental data that demonstrated that human search accuracy could decay with time, we have developed a novel retrial queuing model of an operator that provides the operator with an additional choice by allowing the operator to requeue challenging targets.

A human-in-the-loop experiment was performed under different requeue conditions, which showed that the use of requeues increases the overall probability of detection and the operator confidence but could decrease the fraction of targets found. Furthermore, the additional use of requeues increased operator workload as measured by utilization, although this was shown to be within acceptable standards based on previous research in human supervisory control.

These observations open up an interesting area of work that should seek to understand the value of information of an image, and under what circumstances, an operator may require additional imagery to reach a conclusion. Another important conclusion of this paper is that it has shown that supplying the operators with one additional choice of requeuing, rather than constraining them to a forced choice context, improves accuracy and confidence. This has important ramifications, not only for the external validity of the 2-AC models but also for practical considerations in actual missions, in designing DSSs that can provide additional flexibility to stressed operators.

Future work will include developing "optimal" relook policies, understanding that, in practicality, generating satisficing parameters is more realistic since optimality may be difficult to quantify in dynamic uncertain command and control settings [47]. These policies will also be evaluated in tasks where the target is absent, leading to a richer set of operator models. Moreover, additional work is needed to more fully understand the information processing ramifications of relooks as it is not clear whether the success of the relook mechanism is due to scene complexity, a possible attention filtering bias, or that operators had more confidence knowing that they had such a tool available. Such understanding could possibly lead to identification of images in advance that could cause operator difficulty, possibly allowing them to be inserted in the queue at a more opportune time.

An additional consideration would be to quantify what kind of additional imagery information would be desired by an operator to increase the likelihood of detection in the event of a relook. Finally, a tighter coupling between the role of requeuing and the mission parameters needs to be made. For example, it will be beneficial to understand precisely what the role of vehicle routing is for the purposes of aiding the relook tasks (e.g., with different path planners), as well as the number and heterogeneity of UAVs.

References

[1] C. Drew, Military Is Awash in Data From Drones. [Online]. Available: http://www.nytimes.com/2010/01/11/business/11drone.html

[2] J. Clapper, J. Young, J. Cartwright, and J. Grimes, *Unmanned Systems Roadmap 2007–2032*. Washington DC: Office of the Secretary of Defense, U.S. DOD, 2007.

[3] M. Hoffman, New Reaper Sensors Offer a Bigger Picture. [Online]. Available: http://www.airforcetimes.com/news/2009/02/airforce_WAAS_021609/

[4] M. Cummings and P. Mitchell, "Predicting controller capacity in remote supervision of multiple unmanned vehicles," *IEEE Trans. Syst., Man, Cybern. A, Syst., Humans*, vol. 38, no. 2, pp. 451–460, Mar. 2008.

[5] D. K. Schmidt, "A queuing analysis of the air traffic controller's workload," *IEEE Trans. Syst., Man, Cybern.*, vol. SMC-8, no. 6, pp. 492–498, Jun. 1978.

[6] C. Nehme, "Modeling human supervisory control in heterogeneous unmanned vehicle systems," Ph.D. thesis, MIT, Cambridge, MA, 2008.

[7] R. Bogacz, E. Brown, J. Moehlis, P. Holmes, and J. D. Cohen, "The physics of optimal decision making: A formal analysis of models of performance in two-alternative forced choice tasks," *Psychol. Rev.*, vol. 113, no. 4, pp. 700–765, Oct. 2006.

[8] J. R. Busemeyer and A. Diederich, "Survey of decision field theory," *Math. Soc. Sci.*, vol. 43, no. 3, pp. 345–370, Jul. 2002.

[9] J. R. Busemeyer and J. T. Townsend, "Fundamental derivations from decision field theory," *Math. Soc. Sci.*, vol. 23, no. 3, pp. 255–282, Jun. 1992.

[10] A. Diederich and J. R. Busemeyer, "Simple matrix methods for analyzing diffusion models of choice probability, choice response time, and simple response time," *J. Math. Psychol.*, vol. 47, no. 3, pp. 304–322, Jun. 2003.

[11] R. J. Herrnstein, "Rational choice theory: Necessary but not sufficient," *Amer. Psychol.*, vol. 45, no. 3, pp. 356–367, Mar. 1990.

[12] J. Wolfe, T. S. Horowitz, and N. M. Kenner, "Rare items often missed in visual searches," *Nature*, vol. 435, no. 7041, pp. 439–440, May 2005.

[13] J. Wolfe, "Guided search 2.0: A revised model of visual search," *Psychonomic Bull. Rev.*, vol. 1, no. 2, pp. 202–238, Jan. 1994.

[14] T. Horowitz and J. Wolfe, "Visual search has no memory," *Nature*, vol. 394, no. 6693, pp. 575–577, Aug. 1998.

[15] J. Wolfe, "What can 1 million trials tell us about visual search?" *Psychol. Sci.*, vol. 9, no. 1, pp. 33–39, Jan. 1998.

[16] C. Drury and M. R. Clement, "The effect of area, density, and number of background characters on visual search," *Hum. Factors, J. Hum. Factors Ergonom. Soc.*, vol. 20, no. 5, pp. 597–602, Oct. 1978.

[17] S. K. Hong and C. G. Drury, "Sensitivity and validity of visual search models for multiple targets," *Theor. Issues Ergonom. Sci.*, vol. 3, no. 1, pp. 85–110, May 2002.

[18] L. Itti, C. Gold, and C. Koch, "Visual attention and target detection in cluttered natural scenes," *Opt. Eng.*, vol. 40, no. 9, pp. 1784–1793, Sep. 2001.

[19] M. Chun and M. Potter, "A two-stage model for multiple target detection in rapid serial visual presentation," *J. Exp. Psychol., Hum. Perception Perform.*, vol. 21, no. 1, pp. 109–127, Feb. 1995.

[20] J. Irvine, "Assessing target search performance: The free-response operator characteristic model," *Opt. Eng.*, vol. 43, no. 12, pp. 2926–2934, Dec. 2004.

[21] M. Peterson, A. Kramer, R. Wang, D. Irwin, and J. McCarley, "Visual search has memory," *Psychol. Sci.*, vol. 12, no. 4, pp. 287–292, Jul. 2001.

[22] J. McCarley, A. Kramer, C. Wickens, E. Vidoni, and W. Boot, "Visual skills in airport-security screening," *Psychol. Sci.*, vol. 15, no. 5, pp. 302–306, May 2004.

[23] D. Bertsekas, *Dynamic Programming and Optimal Control*. Belmont, MA: Athena Scientific, 2005.

[24] M. Pachter, P. R. Chandler, and S. Darbha, "Optimal sequential inspection," in *Proc. IEEE Conf. Decision Control*, 2006, pp. 5930–5934.

[25] R. Swensson, "The elusive tradeoff: Speed vs accuracy in visual discrimination tasks," *Perception Psychophys.*, vol. 12, no. 1-A, pp. 16–32, Jan. 1972.

[26] D. Davies and R. Parasuraman, *The Psychology of Vigilance*. London, U.K.: Academic, 1982.

[27] G. I. Falin and J. G. C. Templeton, *Retrial Queues*. London, U.K.: Chapman & Hall, 1990.

[28] J. Artalejo and A. Gomez-Corral, *Retrial Queueing Systems*. London, U.K.: Springer-Verlag, 2008.

[29] B. D. Choi and K. K. Park, "The M/G/1 retrial queue with Bernoulli schedule," *Queueing Syst.*, vol. 7, no. 2, pp. 219–227, Jun. 1990.

[30] N. Sherman and J. Kharoufeh, "An M/M/1 retrial queue with unreliable server," *Oper. Res. Lett.*, vol. 34, no. 6, pp. 697–705, Nov. 2006.

[31] L. Bertuccelli, N. Pellegrino, and M. Cummings, "Choice modeling of relooks tasks in UAV search missions," in *Proc. IEEE Amer. Control Conf.*, 2010, pp. 2410–2415.

[32] A. Diederich, "Dynamic stochastic models for decision making under time constraints," *J. Math. Psychol.*, vol. 41, no. 3, pp. 260–274, Sep. 1997.

[33] C. Gardiner, *Handbook of Stochastic Methods for Physics, Chemistry, and the Natural Sciences*. New York: Springer-Verlag, 1985.
[34] P. Verghese, "Visual search and attention: A signal detection theory approach," *Neuron*, vol. 31, no. 4, pp. 523–535, Aug. 2001.
[35] M. King, G. Stanley, and G. Burrows, "Visual search processes in camouflage detection," *Hum. Factors, J. Hum. Factors Ergonom. Soc.*, vol. 26, no. 2, pp. 223–234, Apr. 1984.
[36] G. Waldman, J. Wootton, and G. Hobson, "Visual detection with search: An empirical model," *IEEE Trans. Syst., Man, Cybern.*, vol. 21, no. 3, pp. 596–606, May/Jun. 1991.
[37] G. Sperling and M. Melchner, "The attention operating characteristic: Examples from visual search," *Science*, vol. 202, no. 4365, pp. 315–318, Oct. 1978.
[38] L. Huang and H. Pashler, "Attention capacity and task difficulty in visual search," *Cognition*, vol. 94, no. 3, pp. B101–B111, Jan. 2005.
[39] J. Wolfe, E. Palmer, and T. Horowitz, "Reaction time distributions constrain models of visual search," *Vis. Res.*, vol. 50, no. 14, pp. 1304–1311, Jun. 2010.
[40] E. Palmer, T. Horowitz, A. Torralba, and J. Wolfe, "What are the shapes of response time distributions in visual search?" *J. Exp. Psychol., Hum. Perception Perform.*, vol. 37, no.1, pp. 58–71, Feb. 2011.
[41] M. Mouloua and R. Parasuraman, *Automation and Human Performance: Theory and Applications*. Hillsdale, NJ: Lawrence Erlbaum Associates, 1996, ch. 9.
[42] K. Savla, T. Temple, and E. Frazzoli, "Human-in-the-loop vehicle routing policies for dynamic environments," in *Proc. IEEE Conf. Decision Control*, 2008, pp. 1145–1150.
[43] A. Girard, S. Dharba, M. Pachter, and P. R. Chandler, "Stochastic dynamic programming for uncertainty handling in UAV operations," in *Proc. IEEE Amer. Control Conf.*, 2007, pp. 1079–1084.
[44] W. B. Rouse, *Systems Engineering Models of Human–Machine Interaction*. New York: North Holland, 1983.
[45] M. Cummings and S. Guerlain, "Developing operator capacity estimates for supervisory control of autonomous vehicles," *Hum. Factors*, vol. 49, no. 1, pp. 1–15, Feb. 2007.
[46] M. Cummings, A. Clare, and C. Hart, "The role of human–automation consensus in multiple unmanned vehicle scheduling," *Hum. Factors, J. Hum. Factors Ergonom. Soc.*, vol. 52, no. 1, pp. 17–27, Feb. 2010.
[47] W. C. Stirling, *Satisficing Games and Decision Making: With Applications to Engineering and Computer Science*. New York: Cambridge Univ. Press, 2003.

Luca F. Bertuccelli (M'04) received the B.S. degree in aeronautical and astronautical engineering from Purdue University, West Lafayette, IN, in 2002 and the M.S. and Ph.D. degrees in aeronautics and astronautics from Massachusetts Institute of Technology (MIT), Cambridge, in 2004 and 2008, respectively. His Ph.D. thesis addressed robustness to modeling errors in Markov decision processes (MDPs) with applications to unmanned aerial vehicle systems and presented a new sampling method for solving robust MDPs for aerospace applications.

He was a Postdoctoral Associate with MIT in the area of human supervisory control. He joined the United Technologies Research Center, East Hartford, CT, as a Senior Research Engineer in August 2011. His current research interests are in robust planning, integrated human–machine systems, and distributed decision making.

Mary L. Cummings (SM'03) received the B.S. degree in mathematics from the U.S. Naval Academy, Annapolis, MD, in 1988, the M.S. degree in space systems engineering from the Naval Postgraduate School, Monterey, CA, in 1994, and the Ph.D. degree in systems engineering from the University of Virginia, Charlottesville, in 2003.

As a Naval Officer and a Military Pilot from 1988 to 1999, she was one of the Navy's first female fighter pilots. She is currently an Associate Professor with the Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, Cambridge. Her previous teaching experience includes instructing for the U.S. Navy with the Pennsylvania State University, University Park, and as an Assistant Professor with the Engineering Fundamentals Division, Virginia Polytechnic Institute and State University, Blacksburg. Her research interests include human supervisory control, human–automation interaction, human systems engineering, and the ethical and social impact of technology.