Considering Human's Non-Deterministic Behavior and his Availability State When Designing a Collaborative Human-Robots System

Thibault Gateau[1], Caroline P. Carvalho Chanel[1], Mai-Huy Le[1] and Frédéric Dehais[1]

Abstract— The objective of this study is to design a human-robots system that takes into account the non-deterministic nature of the human operator's behavior. Such a system is implemented in a proof of concept scenario relying on a (MO)MDP decision framework that takes advantage of an eye-tracker device to estimate the cognitive availability of the human operator, and, some human operator's inputs to deduce where he is focusing his attention. An experiment was conducted with ten participants interacting with a team of autonomous vehicles in a Search & Rescue scenario. Our results demonstrate the advantages of considering the cognitive availability of a human operator in such a complex context and also the interest of using such a decisional framework that can formally integrate the non-deterministic outcomes which model the human behavior.

I. INTRODUCTION

Autonomous Air Vehicles (AAVs) systems are increasingly becoming more "autonomous" in achieving tasks in a variety of operational contexts (e.g. border security, inspection of hazardous areas, military operations). Despite these technical advances, the human is "still vital" [1] to operate these automated systems, as stated by regulation authorities. Generally, the role of the human operator is to supervise the achievement of the mission, to take over during failure, and to make appropriate tactical and ethical decisions [2].

However, AAVs safety reports state that the human factors issues were involved in 80% of mishaps [3]. A deep analysis of these accidents reveals that they usually involve human-automated systems interactions deficiencies rather than human errors *per se*. In fact, the developments of AAVs were generally achieved without taking into account the integration of the human operators [4]: the human operator is commonly considered as a providential agent to take over when the artificial systems fail to adapt [5]. The complexity of theses systems reduces the human operator's abilities to take immediate and appropriate actions [6], especially after long periods of inactivity while monitoring AAVs by a user interface [7], [8]. In addition, poor user interface design and careless implementation of authority sharing confuse the human operator when mental workload exceeds human capacity [9], resulting in poor human-system performance [10]. One has to consider that this topic is getting critical as research efforts are currently oriented on the control of multiple AAVs by one (or few) human operator(s) [11], [12].

[1]Université Fédérale de Toulouse Midi-Pyrennées, ISAE-SUPAERO, Institut Supérieur de l'Aéronautique et de l'Espace, 10, Av Edourd Belin - BP 54032 - 31055 Toulouse Cedex 4 - FRANCE `name.surname@isae-supaero.fr`

A promising avenue to deal with these issues is to consider that robot and human abilities are complementary and are likely to provide better performance when joined efficiently than when used separately. This approach, known as mixed-initiative [13] defines the role of the human and of the artificial agents according to their recognized skills. A key issue in designing such mixed-initiative interactions is to implement a decision system. This latter defines the role and the authority of human and artificial agents, while estimating the capabilities of the human, the robotic agent's status, and the achievement of the mission. Therefore, a first step is to implement methods and tools that can assess the "cognitive state" of the human operator. Many researches have been achieved in order to implement physiological or neuro-physiological based inference systems to deduce cognitive limitations (eg. [14], [15]) in order to dynamically adapt human-robot interactions [10]. Despite these previous works, more research is still needed to improve their reliability due to intra and inter-individual differences. For instance, eye-tracking techniques appear to be suitable in assessing the human operator's cognitive state in the context of a human-AAVs interaction. Desktop eye-trackers can be easily set up below the user's interface and allow the measurement of the operator's eye movement. Many metrics can be collected in real time (for a review see [16]) to estimate cognitive impairment such as attentional tunneling [17], attentional confusion [18] or drowsiness [19].

The main idea of considering human operator's cognitive state is to merge such measurements with the states of the robots as well as the level of achievement of the mission. A decisional policy needs to be set up to manage the dynamics of the human-AAVs interactions, and possibly the AAVs' behavior. On one hand, decision-making systems can be governed by a plan [20] to handle human-robot teaming scenarios using classical planning approaches. The drawback of such solutions is the impossibility to consider uncertainties on actions, partial observable states or potentially non-deterministic behavior of the human operator. On the other hand, the decision-making system can be driven by a policy resulting from the resolution of a (Partially Observable) Markov Decision Process ((PO)MDP) [21], where the two types of agents, human or artificial, may face unexpected situations during the mission or having a partial observation of the state. These outcomes can be modeled as probabilistic effects of actions and as a probabilistic observation function respectively. In this sense, a framework has been proposed [22], which is able to adapt itself to the user's intention, getting feedback from the user in terms of satisfaction. A

different way to manage interactions using POMDPs has been studied for assisting persons with dementia during hand-washing [23]. Nikolaidis et al. (2015) [24] proposed a MOMDP framework where the human is seen as a partially observable state variable in a joint action context involving a human and a robot that automatically learns the human's specific behavior. In another example [25], a decomposition - between the state vector of the robot considered as fully observable while the operator's cognitive state is considered as partially observable - is explored to model the mixed-initiative framework as a Mixed Observability Markov Decision Process (MOMDP) [26]. However, this relevant work does not evaluate experimentally the human operator's behavior once integrated in such a system and the study only provides simulations results.

In this paper, we propose to model and to experimentally evaluate a scenario involving a team composed of AAVs and a human operator. The decisional policy is driven by a MOMDP model that aims to take into account the human's inputs and measures about his cognitive state in order to adapt the AAV's decision autonomy and so optimize the performance of the team. To the best of the authors' knowledge, such experimental evaluation has never been addressed yet. The main contribution of this work consists in:

(1) Including the human's cognitive state in the mission planning model, i.e considering the human's cognitive availability in the model to obtain a policy which drives the mixed-initiative process. The human cognitive's availability is derived from basic eye-tracking metrics and is also used in the on-line state estimation process, which provides appropriate actions for the whole system.

(2) Handling imperfections of the state transition function in the planning model, using an adapted decision model framework. The human operator's behavior is not deterministic. For instance, the reaction time can depend on the person solicited and on the context in which the person evolved. We propose therefore to take advantage of a decision framework that handles such uncertainties as probabilistic actions effects and as probabilistic observation function, and to apply such a decision framework to the system (AAVs and human operator). To this end, the (MO)MDP represents an elegant framework to manage such a decision problem.

(3) Building a proof of concept scenario tested in a well controlled experiment based on a collaborative mission including one human operator and AAVs.

The paper is organized as follows: first we present the materials and methods used in our study case. Next, the collaborative human-AAVs mission problem is formalized based on a MOMDP framework. After, the experimental protocol, metrics and the results obtained are presented. Finally we conclude this paper by discussing future works.

II. EXPERIMENTAL SCENARIO

Humans are generally confronted to concurrent tasks while performing system supervision. In this sense, the human operator was confronted to a dual-task paradigm in the scenario proposed. The first task, defined as *Search & Rescue*

task, consisted of a classical search and rescue mission [27]. The human operator had to collaborate with autonomous artificial agents to perform target identifications by means of an user interface as shown on the left side of Figure 1. The second task, defined as *Short-Term Item Memorization* task, consisted of a series of digits that the human operator had to memorize and report via the user interface as shown on the right side of Figure 1.

It was expected that the volunteers to have a lower performance when achieving both tasks than when achieving each of them separately.

A. Search & Rescue Task

Targets appeared on the simulation screen (the upper part of the AAVs monitoring interface on Fig. 1). A starting point base was defined for the team of AAVs that must: (1) visit target location; (2) take pictures of the target; (3) identify the target; (4) bring the target back to the base if it was a human or a material; (5) avoid bringing back targets that are neither human nor material.

Drones could request help from the human operator for target identification, according to the decisional policy computed (Cf. Section III-B.8). A request to the operator was materialized using a yellow bounding box visual effect on the target, in the simulation screen. For the human operator, the identification task was emulated by a dot counting, depending on the target's real nature (the lower part of *Search & Rescue* task interface - Fig. 1). The identification picture was available for a limited time on the screen, depending on the decisional policy, adding some time-pressure to the task [11], and also to avoid instant target identification, subitizing behaviors were refrained by using more than six dots [28].

B. Short-term Item Memorization Task

A key executive function when operating complex systems is the working memory [14]. In this sense, we chose to design a well know task, the digit span task, to measure Working Memory (WM) storage and update capacity [29]. The volunteers were instructed to retain digits displayed on half of a 21 inches gray screen (on the right part, resolution
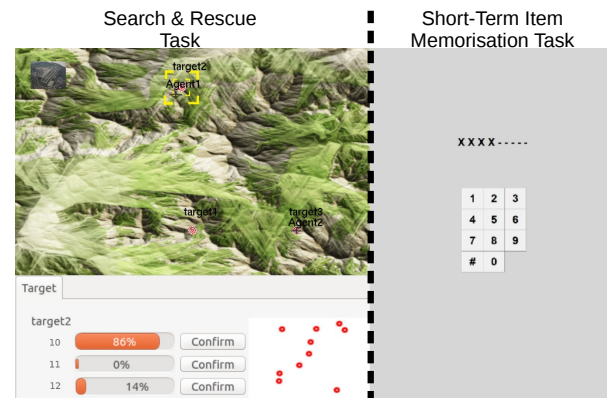

Fig. 1. Screen-shot of the participant's view of the dual task experimental setup.

1680x1050), *via* a script written in Python 2.7.3. Volunteers were asked to use a visual keypad to enter the sequence of seven digits using a mouse cursor, 500 ms after the last digit was presented. A visual keypad was preferred to refrain participant to look outside the screen while entering his response.

A series of digits (i.e. a trial) started with a 13 second-cross fixation. A variable stimulus onset asynchrony (SOA) of 1, 2, or 3 seconds extended the presentation of the fixation cross. Each digit was presented during 500 ms at a rate of 500 ms (i.e. an inter-stimulus interval of 500 ms). A trial ended after a 8 seconds response window. Therefore, a trial lasted 30 seconds in average. The complete task was composed of 20 trials and was designed to last in average 10 minutes.

The difficulty of such a task depends on the number of digits. This way, the difficulty was set to seven digits. In fact, the human WM is fundamentally limited to store seven (more or less two) items [29]. Hence, volunteers were expected to get sufficiently involved into this WM task, without actually reaching mental overload when achieving this task alone [9].

III. COLLABORATIVE PROBLEM FORMALIZATION

A. MOMDP model formalism

The MOMDP model [26] is an extension of the POMDP model [21], which explores the fact that certain state variables are fully observable. Formally, a MOMDP is a tuple $(\mathcal{X}, \mathcal{Y}, A, \Omega, T_{\mathcal{X}}, T_{\mathcal{Y}}, \Omega, R, \gamma, b_0)$, where: $\mathcal{X}$ (resp. $\mathcal{Y}$) is the set of fully (resp. partially) observable state variables; A is the set of actions; Ω is the set of observations; $T_{\mathcal{X}}$ (resp. $T_{\mathcal{Y}}$) is the conditional probability transition function such that, $T_{\mathcal{X}}(x, a, x', y) = p(x'|x, y, a), \forall (x, a, x', y) \in \mathcal{X} \times A \times \mathcal{X} \times \mathcal{Y}$ (resp. $T_{\mathcal{Y}}(x, y, a, x', y') = p(y'|x, y, a, x')$, $\forall (x, y, a, x', y') \in \mathcal{X} \times \mathcal{Y} \times A \times \mathcal{X} \times \mathcal{Y}$); O is the conditional probability observation function such that, $O(o, a, x', y') = p(o|x', y', a), \forall (o, a, x', y') \in \Omega \times A \times \mathcal{X} \times \mathcal{Y}$; $R : \mathcal{X} \times \mathcal{Y} \times A \to \mathbb{R}$ is the reward function defined for a state-action pair; $\gamma \in [0, 1[$ is the discount factor; and, $b_0 = (x_0, b_{\mathcal{Y}_0}) \in \mathcal{B}$ is the initial probability distribution over states, where $\mathcal{B}$ is complete belief state space.

As the artificial decider agent does not have access to the true state of the world, it maintains a probability distribution over the partial observable state variables (initialized by $b_0 = (x_0, b_{\mathcal{Y}_0})$). And, after each action a executed, each next fully observable state x' reached and each observation o perceived, the probability distribution must be updated using the Bayes' rule (η being a normalization constant):

$$b_{\mathcal{Y}}^{o,a,x'}(y') = \eta \sum_{y' \in \mathcal{Y}} p(o|y', x', a)p(y'|x, y, a, x')p(x'|x, y, a)b_{\mathcal{Y}}(y)$$

Note that the belief state b is a tuple formed by the couple $(x, b_{\mathcal{Y}})$, with $b_{\mathcal{Y}} \in \mathcal{B}_{\mathcal{Y}}(x)$. $\mathcal{B}_{\mathcal{Y}}(x)$ is a belief state space over $\mathcal{Y}$ conditioned by x: $\mathcal{B}_{\mathcal{Y}}(x) = \{(x, b_{\mathcal{Y}}), b_{\mathcal{Y}} \in \mathcal{B}_{\mathcal{Y}}\}$. Notice that $\mathcal{B}_{\mathcal{Y}}(x)$ is a sub-space of the complete belief state space $\mathcal{B}$, such that $\mathcal{B} = \bigcup_{x \in \mathcal{X}} \mathcal{B}_{\mathcal{Y}}(x)$.

The MOMDP solving can be seen as an optimization problem for finding *a set of policies* $\pi_x : \mathcal{B}_{\mathcal{Y}} \to A$, which maximizes the criterion:

$$\pi_x^* \leftarrow \arg\max_{\pi_x \in \Pi} E_{\pi_x} \left[\sum_{t=0}^{\infty} \gamma^t r((x_t, b_{\mathcal{Y}}^t), \pi((x_t, b_{\mathcal{Y}}^t))) \middle| b_0 \right].$$

Such a criterion allows the definition of a value function and the use of dynamic programming to solve the problem. Then, the resulting *set of policies* $\pi_x : \mathcal{B}_{\mathcal{Y}} \to A$, that can be extracted from the optimal value function, defines an optimal action which maximizes the expected gain for each reachable belief states $(x, b_{\mathcal{Y}})$, considering that optimal actions would be performed for all possible successive belief states. For more details about MOMDP resolution, please refer to [26].

B. MOMDP model for the collaborative mission

The collaborative mission consists in identifying targets, that can be either human, material, or none, and to bring to the base only relevant targets (Cf. Section II-A). The drones are able to sense the type of targets, but they have a less reliable probability of good classification than the one of the human operator. The human operator plays cooperatively with the drones in order to help them with a sensing task.

1) States, actions and observations: The MOMDP state space is composed by the cross-product between a set of state variables: the mission state of each drone i ($drone_i$), the input of the operator (*operator-input*), the *availability* of the human operator and the type of each target j (*target_j-type*). For the following, let $n_d \in \mathbb{N}^*$ be the number of drones, $n_t \in \mathbb{N}^*$ be the number of targets, $i \in [\![1; n_d]\!]$ and $j \in [\![1; n_t]\!]$.

Possible values for each discrete state variable are:

- $drone_i$: { *going-to-target₁... going-to-target_{n_t}, at-target₁... at-target_{n_t}, going-to-base, base*}, $\forall i \in [\![1; n_d]\!]$;
- *operator-input* : {*at-target₁ ... at-target_{n_t}, other-task, none*};
- *availability*: {*OK, KO*};
- *target_j-type* : {*human, material, none*}, $\forall j \in [\![1; n_t]\!]$.

We consider that the drones' position, the operator's input and the availability are fully observable state variables; and, the target's type is a partially observable state variable.

Therefore we have some discrete observation variables. The drones and the operator can observe the target's type. So, we have:

- $o\text{-}drone_i$: {*ohuman, omaterial, onone*}, $\forall i \in [\![1; nd]\!]$;
- $o\text{-}operator$: {*ohuman, omaterial, onone*};

As actions are defined in a centralized way, the joint action set is composed by the cross product between drones' individual actions and user's interface possible actions. The user interface (i.e. ground station) action is used to relay a drone's requests to the human operator to look at a target (e.g. launching a yellow bounding box) in order to bring more reliable measures about the target type. So, we have:

- $a\text{-}drone_i$: {*goto-target₁...goto-target_{nt}, sample, check, base*}, $\forall i \in [\![1; nd]\!]$;
- $a\text{-}operator\text{-}request$: {*show-target₁...show-target_{n_t}, show-other-task*}.

The transition function defines the dynamic of the model for each agent. We choose to define a clock of 1 second to trigger the model (ask for a new action). Therefore, all transition probabilities related to actions' effects are defined hereafter based on this clock.

2) Effects of drones' actions on drones' states: the action *going-to-target$_j$* moves the drone$_i$ to the *going-to-target$_j$* state with a probability of 1.0, except if the drone$_i$ is on the *going-to-target$_j$* state. In this case, it moves the drone$_i$ to the *at-target$_j$* state with a probability of 0.2, and remains on the *going-to-target$_j$* state with 0.8. This helps us to indirectly model the drones' travel time for a movement that lasts approximately 5 seconds in our application. The action *check* does not change the state of drone$_i$. It consists in sensing information about the target's type that the drone$_i$ is overflying. The action *sample* samples the considered target and changes the drone$_i$'s state to *going-to-base* state with a probability of 1.0. The action *go-to-base* moves the drone$_i$ to the *going-to-base* state with a probability of 1.0, except if the drone$_i$ is already on the *going-to-base* state. In this case it moves the drone$_i$ to the *base* state with a probability of 0.20, and remains in *going-to-base* state with 0.8.

3) Effects of actions related to operator's request: the action *show-target$_j$* depends on the *availability* state of the human operator. The probability effect of this action depends on the time taken by the human operator to answer the drones' request when *available*. To ask for the operator's attention, the ground station performs a visual effect to emphasize where the operator should look (e.g. by launching a yellow bounding box, see Fig. 1). Previous work has shown that the yellow bounding box visual effect that is used on the *Search & Rescue* task is very efficient, taking around 3 seconds in average to be seen [30]. Based on it, we consider that the operator would answer to the interface and so, changes its state to *at-target$_j$* with a probability of 0.33 if *available*, and would remain on the previous state with the residual probability. The action *show-other-task* has no visual effects on the application. When the operator is considered as *not available* all actions have no effect.

4) Effects of drones' actions on target states: when drones are moving, target states do not change. But, when a drone samples a target, this target can change its type with an uniform probability over types. This allows us to model the fact that new targets can be integrated into the model, appearing anywhere on the application, once the actual one has been sampled. Also, when all the drones are on the base, all targets' types can change with an uniform probability. This trick allows us to exploit longer missions without reinitializing the planning problem.

5) The availability of the operator: the operator must perform the drones' monitoring task in priority. But, he can be requested when he is estimated available, e.g. when he has finished the *other-task* (*Short-Term Item Memorization* task) for instance. The *other-task* has a recurrent time duration of 30 seconds in average, with a 13 seconds cross fixation period which is considered as a break where the human operator could be estimated as *available*. Therefore, the human operator has a probability of 0.08 to switch from *KO* to *OK*, which is a worst-case probability, and vice-versa.

6) Observation function: the observation model is based on the fact that all recognition tasks performed by a human operator are still more reliable than when performed by a

state/obs	Drone (Operator)		
	ohuman	omaterial	onone
human	0.65 (0.95)	0.30 (0.05)	0.05 (0.05)
material	0.20 (0.05)	0.75 (0.95)	0.05 (0.05)
none	0.05 (0.05)	0.05 (0.05)	0.90 (0.95)

TABLE I

DRONES (AND HUMAN) OBSERVATIONS PROBABILITY TABLE.

drone. An informed observation can be done by a drone (resp. the human) only if it is over a target and only if it performs a check action (resp. only when he answers a *show-target* action request). The observation function is summarized in Table I. These probabilities were arbitrarily chosen for this study case.

7) Reward function: all moving actions have the same cost, 0.1 per drone, except if at least one drone is on the base (it is discounted from the cost). The *check* action has the cost of 5 per drone. We consider that the payoff related to a *sample* action is superior for human targets than for material ones (human 1000, material 300). If the drone samples a non relevant target it costs 10000 per target sampled. Note that if two drones sample the same target the payoff remains the same. It prevents multiple drones to sample the same target. The action *sample* is penalized if it is performed elsewhere than while flying a target. If a *show-target* request action is performed, it costs 2.5 if the operator is *available*, and 20 if he is *not available*.

8) Off-line policy generation: for this study case, and based on the parametric model before presented, we have generated the MOMDP model considering 2 drones and 3 targets. The model obtained was solved using the SARSOP algorithm [26], with $\epsilon = 0.1$ and with a memory limited to 6.5Gb (a computer with 8Gb RAM). The computed policy was then executed on the experimental platform described hereafter to handle the collaborative interaction between human and artificial agents.

IV. MATERIALS AND METHODS

A. Volunteers

Ten male volunteers (mean age: 33.0, STD= 8.1) completed the experiment. They have normal or corrected-to-normal vision, normal hearing, and no psychiatric disorders. After providing written informed consent, they were instructed to complete the experimental scenario (Cf. Section II).

B. Global System Architecture

The global system architecture is presented in Fig. 2. The interaction between the operator and the AAVs is ensured by the ATMOSPHEr[1] interface [31]. This interface can be seen as part of a ground station which is equipped with an eye-tracker sensor.

---

[1]A Tiny Multi-agent system Oriented Simulator Platform for HMI Experiments - https://sourceforge.isae.fr/projects/atmospher/
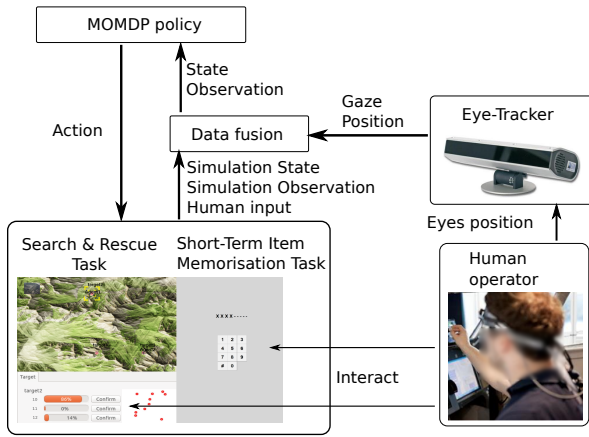
Fig. 2. The complete system architecture. The human operator is able to interact with the *Search & Rescue* task and the *Short-Term Item Memorization* task interfaces, while an eye-tracker is monitoring his eyes position. The ATMOSPHEr interface sends to the `Data Fusion` module the current drones' state and observations, and the human operator's input (sensing task) with a frequency of 1Hz (the same clock used on the MOMDP belief update and action request too). The `Data Fusion` module sends to the MOMDP policy module the compiled couple (State, Observation). The MOMDP module updates the belief state and sends the next Action to the ATMOSPHEr interface.

1) Eye-tracking Device: An eye-tracker (SMI RED250) was used to collect the human operator's eye gaze. Volunteers were asked to stay at a distance of around 60 cm from the screen. The calibration tolerance was set to be horizontally and vertically under $0.5°$. The refreshment rate of the eye-tracker was set to 60Hz. We used SMI iView X^{TM} Software Development Kit[2] to retrieve the volunteer's gaze position on the screen, and to export it through a Python TCP/IP interface that communicates with the `Data Fusion` module (Fig. 2). We used the participants' eye gaze as a straightforward indicator of their cognitive availability. In a nutshell, the discrete state variable *availability* (Cf. Section III-B.1) for the request of the decisional policy was set to *OK* when the participant gaze was on the left part of the screen (looking at *Search & Rescue* task), and *KO* otherwise (looking at *Short-Term Item Memorization* task or outside the screen, or if the gaze signal was lost).

C. Metrics

1) Performance evaluation on Search & Rescue task: We evaluated the operator's *Search & Rescue* task performance with he ratio of the number of answered requests over the total number of requests asked by the policy. We were interested in how much the human operator was engaged on this task, that was defined as primary. We also logged the expected rewards of the human-robots system during experiments (MOMDP score given by the SARSOP policy simulator) as an indicator of policy execution efficiency (Cf. Section III-B.7).

2) Perfomance evaluation on Short-Term Item Memorization task: Each participant had the same number of digits series during each session. Therefore, we defined the performance criteria of *Short-Term Item Memorization* task

---
[2]lyrawww.uvt.nl/ cenv/dci-lab/smi/iViewX.pdf

---

as the number of digits series that has been fully retrieved, in the correct order.

D. Protocol's Sequence

First, the volunteer was trained on the *Short-Term Item Memorization* task alone (duration: 10 min), and then on the *Search & Rescue* task alone (duration: 10 min). Afterwards, the volunteer passed two times the dual task protocol:

- in one condition, the decisional system was using the eye gaze position measured by the eye-tracking device: we will call it the "with gaze monitoring" (*with GM*) condition (duration: 10 min);
- in the other condition, the decisional system had no information about where the volunteer was looking, and it was considering him as "available": this was the "without gaze monitoring" (*without GM*) condition (duration: 10 min) and was used as a control condition.

The order between the *with GM* condition and the *without GM* condition was counter-balanced between the volunteers. The volunteers had no information about the type of condition they were achieving. At the end of each dual task condition, the volunteer completes a NASA-TLX form [32]. The typical total duration of a subject's session (informed consent approval, the two practice tasks, and the two dual tasks) was about one hour and fifteen minutes.

V. EXPERIMENTAL RESULTS

A. The Search & Rescue task's Scores

The *Search & Rescue* task's scores are summarized in Table II. One can observe that there is a tendency concerning a better request ratio score on *with GM* condition than on *without GM* condition. However, as the volunteers' request ratio score is not normally distributed, non-parametric Wilcoxon rank-sum tests were applied on median values for each condition. These tests reveal that ratio scores are not significantly different between the *with GM* condition, the *without GM* condition and the *Search & Rescue* task achieved alone. Based on this, we can not formally advance any conclusion with respect to this score. It is also worth noting that the variance was the largest in the *without GM* condition suggesting that the volunteers had more contrasted behavior regarding task prioritization. The strategy in the *with GM* condition proposed by the system led to less variability reflecting a more balanced behavior of human operators.

In the same Table, the average expected rewards were also summarized. Given that these expected rewards were averaged between 10 volunteers, this is obviously not enough to evaluate the performance of the MOMDP policy. Paired t-tests did not reveal any significant difference concerning final expected reward (human-robots system) between the three conditions. Nevertheless, average expected rewards have the same order of magnitude. It means that even when the human operator was not available to help the AAVs (*with GM* condition), the robotic system could achieve the target recognition task alone.

Metric	Request ratio score	Averaged discounted sum of rewards
without GM	0.76 (0.26)	1321.4 (456.5)
with GM	0.84 (0.11)	1210.4 (269.5)
S&R task alone	0.92 (0.03)	1143.4 (469.9)

TABLE II

Search & Rescue TASK'S AVERAGED SCORES (AND STD).

B. Short-Term Item Memorization task's Scores

The *Short-Term Item Memorization* task's scores are summarized in Figure 3. For this scores, parametric two-tailed paired t-tests were applied on mean values for each condition. P-values were corrected for multiple comparisons (Bonferroni-Holm). Paired t-test results revealed that the volunteers perform significantly better on *Short-Term Item Memorization* task during the *with GM* condition than in the *without GM* condition ($p_{corrected}$<0.01) (Fig. 3). They also performed significantly better during the *Short-Term Item Memorization* task alone training session than in the *with GM* condition ($p_{corrected}$<0.001).
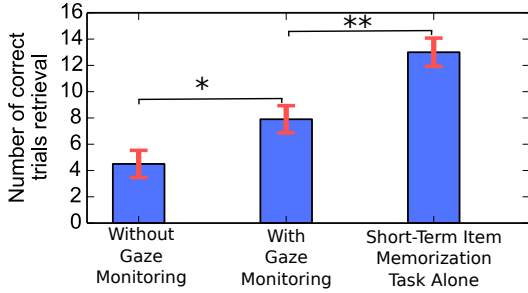

Fig. 3. Mean number of fully correct digits on the three conditions (*without GM*, *with GM* and training on *Short-Term Item Memorization* task alone), on 20 trials. Red bars indicate standard errors and *: $p_{corrected} < 0.01$, **: $p_{corrected} < 0.001$.

C. Volunteers' Subjective Feedback

A two tailed paired t-test was conducted on the NASA-TLX scores and revealed that the volunteers rated the *with GM* condition's workload statistically significantly lower than the *without GM* condition's workload ($p<0.01$). These results show the importance of taking into account a measure of the human self-estimated workload while interacting with robots and performing many tasks.

D. Synthesis

The behavioral measures reveal that the decision system including a cognitive state indicator (*with GM* condition) helps the participant to maintain a good performance on the primary drones monitoring task (*Search & Rescue* task), while improving the score on the secondary WM task (*Short-Term Item Memorization* task). Consistently, the subjective rating confirmed that *with GM* condition reduced the volunteer's workload. Our results confirm previous findings [1] supporting that taking into account the human operator in the design of the system leads to better overall performance. In our context, the overall performance is the performance of the whole system that includes human-robots and all the tasks the human has to perform.

VI. CONCLUSION AND FUTURE WORKS

The objective of this study was to show that the (MO)MDP decision framework is suitable to design a human-robots interactive system. We implemented such a framework to handle the non-deterministic nature of the human operator's cognitive availability.

The results of this study supports the advantages of taking into consideration the cognitive availability of the human operator in such complex contexts and to integrate psychophysiological measurement: volunteers had better objective performances in the double task paradigm when their cognitive availability state was taken into account. The subjective measurements were also in line with these latter results as the volunteers experienced lower mental workload in the *with GM* condition than in the *without GM* one. Taken together these results demonstrate that the human is not a providential agent and that he can not take over or assist the robots at anytime.

The MOMDP formalism is suited to model probabilistic actions' effects, as the human's time to answer a request when solicited, and probabilistic observations as the misclassification rate when identifying targets. Moreover, such a formalism can be used to consider the human cognitive state as a partial observable state (see [24], [25] for instance). More interestingly, using an adapted decision model framework, such as a MOMDP, we can directly integrate confidence weight on the observation of the cognitive state we are obtaining. Though we admit that our metric to derive the cognitive availability was simple, this proof of concept demonstrates that it is easy to introduce uncertainties on observations concerning this state variable while maintaining the same framework. Therefore, we advocate for the suppleness of such a MOMDP formalism, that allows to easily extend the model to consider new state variables and observation variables. This approach paves the way to integrating more complex partial observable human cognitive states, such as mental workload, mind wandering state, and task engagement. Improvements can also be made concerning the model used (Cf. Section III): for example, the generation of the probability table (Cf. Section III-B.6) could be improved by using learned data. If probabilities tables are hard to be learned, one could consider an extension of the MOMDP model, the Qualitative Possibilistic MOMDPs [33], where the possibility theory is applied instead of a probability one.

REFERENCES

[1] R. Parasuraman and C. D. Wickens, "Humans: Still vital after all these years of automation," *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 50, no. 3, pp. 511–520, 2008.

[2] J. Galliott, *Military Robots: Mapping the Moral Landscape*. Ashgate Publishing, Ltd., 2015.

[3] K. W. Williams, "A summary of unmanned aircraft accident/incident data: Human factors implications," DTIC Document, Tech. Rep., 2004.

[4] N. Schurr, P. Scerri, and M. Tambe, "Impact of human advice on agent teams: A preliminary report," in *Workshop on Humans and Multi-Agent Systems at AAMAS*, 2003.

[5] N. Schurr, J. Marecki, and M. Tambe, "Improving adjustable autonomy strategies for time-critical domains," in *Proceedings of The 8th International Conference on Autonomous Agents and Multiagent Systems-Volume 1*. International Foundation for Autonomous Agents and Multiagent Systems, 2009, pp. 353–360.

[6] D. D. Woods and N. B. Sarter, "Learning from automation surprises and" going sour" accidents," *Cognitive engineering in the aviation domain*, pp. 327–353, 2000.

[7] M. Boyer, M. Cummings, L. B. Spence, and E. T. Solovey, "Investigating mental workload changes in a long duration supervisory control task," *Interacting with Computers*, p. iwv012, 2015.

[8] G. Durantin, F. Dehais, and A. Delorme, "Characterization of mind wandering using fNIRS," *Frontiers in systems neuroscience*, vol. 9, 2015.

[9] G. Durantin, J.-F. Gagnon, S. Tremblay, and F. Dehais, "Using near infrared spectroscopy and heart rate variability to detect mental overload," *Behavioural brain research*, vol. 259, pp. 16–23, 2014.

[10] C. Tessier and F. Dehais, "Authority management and conflict solving in human-machine systems," *AerospaceLab-journal*, no. 4, 2012.

[11] M. L. Cummings, S. Bruni, S. Mercier, and P. Mitchell, "Automation architecture for single operator, multiple uav command and control," DTIC Document, Tech. Rep., 2007.

[12] M. Barnes and F. Jentsch, *Human-robot interactions in future military operations*. Ashgate Publishing Company, 2010.

[13] J. Allen, C. I. Guinn, and E. Horvtz, "Mixed-initiative interaction," *IEEE Transactions on Intelligent Systems and their Applications*, vol. 14, no. 5, pp. 14–23, 1999.

[14] T. Gateau, G. Durantin, F. Lancelot, S. Scannella, and F. Dehais, "Real-time state estimation in a flight simulator using fNIRS," *PloS one*, vol. 10, no. 3, 2015.

[15] N. Régis, F. Dehais, E. Rachelson, C. Thooris, S. Pizziol, M. Causse, and C. Tessier, "Formal detection of attentional tunneling in human operator–automation interactions," *IEEE Transaction on Human Machin System*, 2014.

[16] K. Holmqvist, M. Nyström, R. Andersson, R. Dewhurst, H. Jarodzka, and J. Van de Weijer, *Eye tracking: A comprehensive guide to methods and measures*. Oxford University Press, 2011.

[17] N. Regis, F. Dehais, C. Tessier, and J.-F. Gagnon, "Ocular metrics for detecting attentional tunnelling," *Human Factors and Ergonomics Society–Chapter Europe, Toulouse, France*, 2012.

[18] F. Dehais, V. Peysakhovich, S. Scannella, J. Fongue, and T. Gateau, "Automation Surprise in Aviation: Real-Time Solutions," in *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*. ACM, 2015, pp. 2525–2534.

[19] Q. Ji and X. Yang, "Real-time eye, gaze, and face pose tracking for monitoring driver vigilance," *Real-Time Imaging*, vol. 8, no. 5, pp. 357–377, 2002.

[20] K. Talamadupula, G. Briggs, T. Chakraborti, M. Scheutz, and S. Kambhampati, "Coordination in human-robot teams using mental modeling and plan recognition," in *Intelligent Robots and Systems (IROS 2014), 2014 IEEE/RSJ International Conference on*. IEEE, 2014, pp. 2957–2962.

[21] R. Smallwood and E. Sondik, "The optimal control of partially observable Markov processes over a finite horizon," *Operations Research*, pp. 1071–1088, 1973.

[22] T. Taha, J. V. Miró, and G. Dissanayake, "A pomdp framework for modelling human interaction with assistive robots," in *Robotics and Automation (ICRA), 2011 IEEE International Conference on*. IEEE, 2011, pp. 544–549.

[23] J. Hoey, A. Von Bertoldi, P. Poupart, and A. Mihailidis, "Assisting persons with dementia during handwashing using a partially observable Markov decision process," in *Proc. Int. Conf. on Vision Systems*, vol. 65, 2007, p. 66.

[24] S. Nikolaidis, R. Ramakrishnan, K. Gu, and J. Shah, "Efficient model learning from joint-action demonstrations for human-robot collaborative tasks," in *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction*, ser. HRI '15. New York, NY, USA: ACM, 2015, pp. 189–196. [Online]. Available: http://doi.acm.org/10.1145/2696454.2696455

[25] P. E. U. de Souza, C. P. C. Chanel, and F. Dehais, "MOMDP-based target search mission taking into account the human operator's cognitive state," in *27th IEEE International Conference on Tools with Artificial Intelligence (ICTAI)*, 2015.

[26] S. C. Ong, S. W. Png, D. Hsu, and W. S. Lee, "Planning under uncertainty for robotic tasks with mixed observability," *The International Journal of Robotics Research*, vol. 29, no. 8, pp. 1053–1068, 2010.

[27] L. E. Parker, "Distributed intelligence: Overview of the field and its application in multi-robot systems," *Journal of Physical Agents*, vol. 2, 2008.

[28] G. Mandler and B. J. Shebo, "Subitizing: an analysis of its component processes." *Journal of Experimental Psychology: General*, vol. 111, no. 1, p. 1, 1982.

[29] G. A. Miller, "The magical number seven, plus or minus two: some limits on our capacity for processing information." *Psychological review*, vol. 63, no. 2, p. 81, 1956.

[30] J.-P. Imbert, H. M. Hodgetts, R. Parise, F. Vachon, F. Dehais, and S. Tremblay, "Attentional costs and failures in air traffic control notifications," *Ergonomics*, vol. 57, no. 12, pp. 1817–1832, 2014.

[31] J. Collart, T. Gateau, E. Fabre, and C. Tessier, "Human-robot systems facing ethical conflicts: a preliminary experimental protocol," in *Workshops at the Twenty-Ninth AAAI Conference on Artificial Intelligence*, 2015.

[32] S. G. Hart and L. E. Staveland, "Development of NASA-TLX (task load index): Results of empirical and theoretical research," *Human mental workload*, vol. 1, no. 3, pp. 139–183, 1988.

[33] N. Drougard, F. Teichteil-Königsbuch, J. Farges, and D. Dubois, "Qualitative possibilistic mixed-observable mdps," in *Proceedings of the Twenty-Ninth Conference on Uncertainty in Artificial Intelligence, UAI 2013, Bellevue, WA, USA, August 11-15, 2013*, 2013.