Towards human-robot interaction: a framing effect experiment

Paulo E. U. Souza
Caroline P. Carvalho Chanel
and Frédéric Dehais
Institut Supérieur de l'Aéronautique et de l'Espace
Université de Toulouse, 31055 Toulouse Cedex 4
name.surname@isae.fr

Sidney Givigi
Department of Electrical and Computer Engineering
Royal Military College of Canada
PO Box 17000, Station Forces
Kingston, Ontario
Email: Sidney.Givigi@rmc.ca

***Abstract*—Decision making is a critical issue for humans operating unmanned vehicles. However, it is well admitted that many cognitive biases affect human judgments, leading to suboptimal or irrational decisions. The framing effect is a typical cognitive bias causing people to react differently depending on the context, the probability of the outcomes and how the problem is presented (loss vs. gain). There is a need to better understand the effects of these biases in operational contexts to optimize human-robot interactions. We therefore conducted an experiment involving a framing paradigm in a search and rescue mission (earthquake) and in a Mars rock sampling mission. We manipulated the framing (positive vs. negative) and the probability of the outcomes. Our findings revealed that the way the problem was presented (positively or negatively framed) and the emotional commitment (saving lives vs. collecting the good rock) statistically affected the choices made by the human operators.**

I. Introduction

In recent years, there has been an increasing discussion about the merits of totally autonomous robots versus the importance of user control and decision making [1]. On the one hand, the advancement in artificial intelligence decision-making techniques for aerial robots, also known as drones, has significantly increased the number of applications for a team of autonomous agents (for instance, search and rescue missions [2]–[4], autonomous infrastructure inspection [5], or autonomous patrolling systems [6]–[8]). On the other hand, the continuous evolution of *Human-Robot Interaction* (HRI) allows human operators to improve more and more their performances in controlling and deciding. *Mixed-initiative interaction* provides the interface between these two worlds considering that the agents' (human and robot) abilities are complementary and are likely to provide better performance when joined efficiently than when used separately [1], [9], [10].

However, poor user interface design, complexity of automation and high operational pressure can leave the human operator ill-equipped when mental workload exceeds human mental budget [11]. For instance, careless design of authority sharing can lead to human-automation conflicts when the human operator misunderstands the automation behavior [12], [13]. The occurrence of such a situation is critical as long as it may cause "mental confusion" [13] or attentional tunneling (i.e., the human operator is excessively focused on a specific area of a display) [14] causing irrational behavior [15].

Moreover, research suggests that when human decision-makers have to make critical decisions under uncertainty and imperfect information conditions or in situation where they are emotionally involved, like during natural disasters, emergencies, military operations or any other unpredictable and diffuse environment, people are more susceptible to make predictable errors in judgment caused by *cognitive biases* [16]. In order to overmatch these situations and remain adaptive and effective amid a complex and ambiguous environment it is important to understand and deal with these *hard-wired* human processes [17].

This study is part of a broader work in a robust mixed-initiative multi-agent planning, control, and execution framework that maximizes joint performance, based on predefined guidelines, by taking into account the abilities and limitations of each agent (see Fig. 1). Here, "robust" means that the system has satisfactory performance even when reality differs from assumptions and the multi-agent team is composed by a group of drones and a human operator.

The guidelines are statements and recommendations that determine a course of action in consonance with the decisions of the authority in charge of the mission. They can suggest the trade-off criteria for deciding during the execution time, for instance, of a specific mission what is more important: accuracy or speed.

In this framework, each drone obtains its movement strategies by using a game-theoretical approach and should adapt the interaction with the operator (via a graphical user interface - GUI) by framing the situation according to the guidelines given by the mission authorities (see Fig. 1). The human operator is in charge of the "hard decisions", under conditions of uncertainty and with time constraints. The framing, in this context, should explore the human cognitive biases [17] in order to improve the overall systems performance.

For this purpose, a framing-effect experiment was carried out, where the drone team was supervised by a human operator, in order to evaluate the presence of such a framing context. For this, two different scenarios were designed: helping vic-

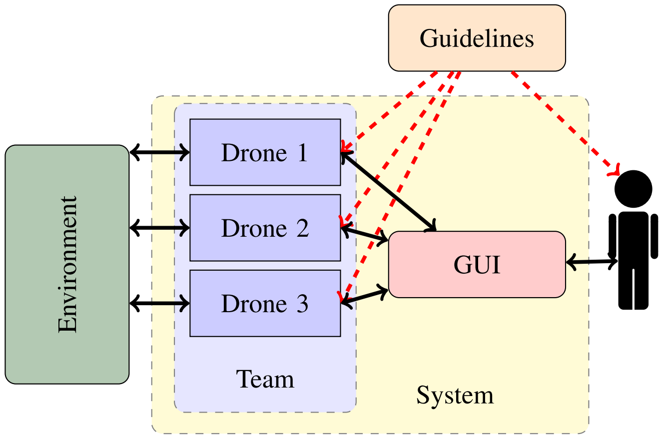

Fig. 1. Framework architecture

tims of an earthquake and sampling rocks on Mars. The final objective is, in future work, to take into account these cognitive biases (1) to adjust the utility functions of the drones and (2) to adapt dynamically human-robot interactions.

This work is organized as follows: Section II presents the theory about *Framing Effect* (FE) and how it can affects the human inferences in a specific situation. In order to verify the FE "power", experiments are carried out as described in Section III. The results of the experiments are presented in Section IV. Finally conclusions and future work are discussed in Section V.

II. Framing effect

According to Descriptive Theories of decision making, the Framing Effect (FE) is the finding that different descriptions of formally identical problems can result in different choices. The core concepts that explain it reside in the combination of beliefs, fears, values, mental models, and so on, which human beings use to perceive a situation. People effectively look through this frame in the way they would look through colored sunglasses. The frame significantly affects how we infer meaning and hence understand the situation [16].

During the past decades, psychologists have been intensely interested in the two main modes of thinking among humans: intuitive and analytical [16]. According to Prospect Theory [18], intuitive thinking operates automatically and quickly, with little or no effort and totally unconscious, allowing people to multitask in a complex world. The crucial benefits of intuitive thinking are that it is time efficient and requires relatively little allocation of mental resources. By generalizing circumstances, it allows us to reduce the complexity of a situation, recognize patterns (real or perceived) and make decisions quickly according to past experiences or the logic of those recognized patterns. However, while this mode of thinking is exceptionally efficient and very often accurate, it makes us more vulnerable to errors.

On the other hand, analytical thinking requires conscious mental effort. It allows us to process information deliberately, consciously contemplate alternatives, debate with others and come to logical and effective conclusions.

Both modes of thinking are constantly active in our minds, but analytical thinking is typically relegated to simply monitoring the cognitive activities and can be called only when necessary. It is activated when we detect an error or when some rule-based reasoning is required [16].

Moreover, research has shown that losses evokes stronger negative feelings than gains [19] and choices are not reality-bound because intuitive thinking is not bound to reality. Reframing is effortful and analytical thinking is normally lazy. Unless there is an obvious reason to do otherwise, most people passively accept decision problems as they are framed and therefore rarely have an opportunity to discover the extent to which their preferences are frame-bound rather than reality-bound.

Kahneman [16] affirms that human beings dispose of a limited budget of attention that can be allocated to activities, and they will fail if try to go beyond their budget. Intense focusing on a task can make people effectively blind, even to stimuli that normally attract attention.

In summary, Prospect Theory shows that emotions can help play a role in decision-making when information is incomplete or too complex, to serve at times as critical rules of thumb.

In [20] a typology is presented to distinguish among three different kinds of framing effects: (1) Risk Choice Framing [21], which involves options differing in level of risk and described in different ways; (2) Attribute Framing, which affects the evaluation of the characteristics of an event or object; and (3) Goal Framing, which affects the persuasiveness of a communication. Attribute Framing seems to be the simplest case of framing, where only a single attribute is the subject of the framing manipulation and the evaluation can be measured by choices between yes or no. Attribute Framing effects are also less likely when dealing with extremes. So, we have chosen to study the Attribute Framing effect influence in a drone operation situation.

III. Experiments

With two scenarios, an experimental protocol was designed to observe the FE influence over the decision of releasing or not a first-aid kit (first scenario) or collecting or not a rock (second scenario). To achieve this, a simulation with a graphical interface (see Fig. 2) common to both scenarios was set up in Python 2.7.11. The left panel has a 3D environment where operators can change the point of view as they wish. During the simulation, three drones (2) with a limited capacity to identify their targets depart from the base (1) to the search zone (3). The control panel at right shows (4) the status of the drones, where the buttons change colors when any of them needs an operator decision; (5) the battery level of each drone; (6) the sectors already visited, here the gray intensity is correlated to the type of search pattern used by the drone; and (7) the number of kits or storage places available for the mission in a given moment.

We suppose, for a such mission, the drones would be driven individually by a decisional framework, which should allow them to take a decision based on an expectation (reward

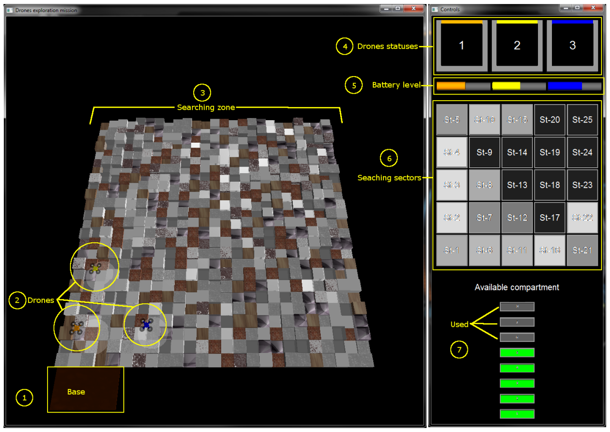

Fig. 2. Operator's interface.

expectation). This expectation would be function of the belief of a drone at the moment of questioning the operator. The belief (percentage, for instance) would define in this case a rational decision.

Next, the two scenarios defined to evaluate the presence of a framing effect are briefly presented.

A. First scenario - Earthquake

In this scenario, an earthquake happened and there are eight victims in the middle of the wreckage. Some of them totally or partially buried under layers of concrete. The operator (subject) use the drones to identify the victims and eight available kits. The guideline is to find and deliver a first-aid kit to the maximum number of victims within a certain time period.

B. Second scenario - Mars rock sampling

Now, the subject is working in a Mars sample return mission where eight different types of rock are searched. The three drones are available for the mission and a capsule with eight storage places is available to return the rocks to Earth. The guideline is to find and collect the maximum number of "good" rocks before the time is up.

C. Research questions

The goal of these experiments is to answer the following research questions:

- *Is the FE present in this context?*
- *Does the emotional commitment influence the FE efficiency?*
- *Has the Time to answer any influence over the FE efficiency?*
- *Do the operator's levels of confidence and satisfaction increase when the framing is aligned to a good choice?*
- *What effects have Positive or Negative framed questions and Probability on the operator's decision?*
- *Is there a choice of Framed questions that could provide a predictable decision regardless the Probability*?

D. Experimental protocol

In order to evaluate the results of this experiment, two explanatory variables have being used: (1) *Text Framing* and (2) the *Probability* that a kit would be useful or not (earthquake scenario) or the target be or not a "good" rock.

The Text Framing is the way how the questions are presented (positive or negative). In the earthquake scenario, for the *Positive frame* a sentence was presented like: "There is 70% of chance that the kit will be **useful**" and for the *Negative frame* it was: "There is 30% of chance that the kit will be **wasted**". In the case of the Mars rock sampling scenario, for the *Positive frame* a sentence was presented such as: "There is 70% of chance of being a **'good' rock**" and: "There is 30% of chance of being a **'bad' rock**", otherwise.

For the *Probability*, four levels of interest were selected: *Low* (from 0.13 to 0.25), *Middle-Low* (from 0.37 to 0.49), *Middle-High* (from 0.51 to 0.63) and *High* (from 0.75 to 0.87). Each level represents a range of 12.5%. Two more levels were introduced with the intention of hiding these levels of interest, they are: *Extreme Low* (from 0.01 to 0.12) and *Extreme High* (from 0.88 to 0.99).

This experiment involved two factors with many levels (2 types of sentences and 4 probability levels). In general, *factorial designs* are most efficient for this situation [22], because in each complete trial of the experiment all possible level combinations of the factors are investigated. In this sense, this factorial design requires 8 runs for each combination (i.e. $2 * 4$) to be tested in each scenario, consequently, each subject must take at least 8 different decisions during the experiment.

Additionally, the confidence and satisfaction levels of the participants about their performance after each mission was checked with a seven-points Likert-type scale. At the end of the experiment they were invited to answer a questionnaire.

E. Execution

Fourteen individuals (28% female, mean age of 30.57 with a standard deviation of 7.63), all volunteers, participated in the experiment. The participants were, unkown to them, randomly split into two groups (one for each scenario). They were not rewarded for the participation.

Each participant randomly executed 10 missions. During the evolution of a given mission, when the drones found something, 10 different types of sentences (2 types of frames * 5 levels of probabilities) were randomly presented and the operator was requested to decide. Every two of the sentences had the same level of *Probability*, one with a positive frame and the other with a negative one. The subjects had 10 seconds to decide between say "YES", i.e, take a positive action (release a kid or collect a rock), or "NO". After this time period, the drone who asked should consider the operator's decision as a "NO".

In this sense, the only action available to the operator was to answer the questions made by the drones, which the participants were oriented to answer *as accurately and quickly as possible*. Note that half of the sentences had less than 50% of probability, then, if an operator chose as strategy to do a positive action (release a kit or sample a rock) only with a probability above that value in a positive frame or below that value in a negative one, fatally he or she would say "YES" only 4 times per mission.

Notice that, it was not possible to know the real result of the mission, i.e., the operator could not know how many victims were helped or "good" rocks were collected during the experiment.

IV. Results

We have collected 1390 observations from 14 subjects. Because we have taken multiple measures per subject, which would violate the *independence assumption* of a linear model (in fact, every person has some idiosyncratic factor that affects all responses from the same subject), a *Linear Mixed Model* was used [23] to deal with this situation. Adding some random effects for subject allows us to resolve this non-independence by assuming a different baseline for each subject.

In this mixed design, we tested our hypotheses comparing the results of the two scenarios using a *Generalized Linear Mixed Model - GLMM* [24]. GLMMs are an extension of linear mixed models to allow response variables from different distributions, such as binary responses ("YES" or "NO").

In this study we have been interested in the relationship between operator's decision ($\boldsymbol{OD}$) and the main explanatory variables: Scenario ($\boldsymbol{S}$), Text framing ($\boldsymbol{TF}$) and Probability ($\boldsymbol{P}$) (see Eq. (1)).

$$\boldsymbol{OD} \sim \boldsymbol{S} + \boldsymbol{TF} + \boldsymbol{P} + \epsilon \tag{1}$$

First of all, the random factors ϵ, that was not possible to control experimentally, were unpacked in two different variables: $\boldsymbol{ID}$ and $\boldsymbol{Seq}$. The first one referred to the assumption of a different intercept for each subject and the second one referred to the sequence of the missions, which were shuffled for each subject. Several *null models* with only these random effects and their interactions were designed and tested (Likelihood Ratio Test - LRT) in order to select the important ones, then, all the others "stochastic" differences have remained in the error term ϵ. The final *null model* is presented in Eq. (2) (LRT:$\chi^2(1) = 10.941, p = 0.0009404$).

$$\boldsymbol{OD} \sim 1 + (1|\boldsymbol{ID}) + (1|\boldsymbol{Seq}) + \epsilon \tag{2}$$

Figure 3 shows a dotplot of the random effect terms. Here is possible to see the effects of each operator and the mission sequence on the decision process as well as their standard errors to help identify how distinct the random effects are from one another. The first plot shows that some subjects (we suppressed their IDs) are more meticulous (negative values) than others in their choices, while the second demonstrates that the subjects used to change their strategies over time, reducing their minimal threshold to do an action. These two effects were informed by 65.2% of the participants in the questionnaire at the end of the experiment.

We started the statistical analysis with a model with all fixed effects available and dropped one by one until all unnecessary terms were removed, for instance: age and gender. In order to check the goodness of fit (GOF) of each model, the *Hosmer and Lemeshow test* [25] was used. And, at the end, the model was checked to make sure that the data were not overdispersed [26].

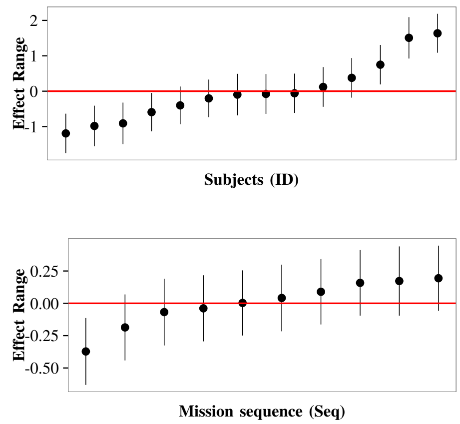

Fig. 3. Random effects by (a) subject and (b) Mission sequence

In order to answer the research questions (cf. Sec. III-C), some hypotheses were made.

A. Hypothesis H1

To answer the first research question, which was *"Is the FE present in this context?"*, we elaborated the following hypothesis:

- $H1_A$: [1] The operators' decision are different, in a similar situation, according to the equivalent version of choice presented, a negative or a positive one.

Here, there is a significant difference ($Positive : Estimate = 0.280, sd = 0.109, z = 2.553, p = 0.0107$) between positive and negative framing in a GLMM analysis (LRT: $\chi^2(2) = 35.997, p = 1.525e-08$). The small $p-value(0.0107)$ indicates that we would reject the *null hypothesis* $H1_0$ in favor of $H1_A$. In this sense, we can confirm the presence of the FE in the experiment.

B. Hypothesis H2

Looking to the next research question: *Does the emotional commitment influence the FE efficiency?*, another hypothesis was defined:

- $H2_A$: The *positive frame* will be more effective in situations of emotional commitment (first scenario).

Results of a GLMM ($LRT : \chi^2(4) = 47.043, p = 1.494e-09$) show that there is a significant difference between the two scenarios ($Mars : Estimate = -0.526, sd = 0.246, z = -2.133, p = 0.032$). The negative estimated value (-0.526) refers to the *Mars* scenario and suggests that people are willing to take more risks when they are emotionally involved (*Earthquake* scenario). Figure 4 shows the influences of *Text framing* and *Scenario* variables in the operators' decisions. The difference of behavior between the two scenarios demonstrates that the participants appeared to take more risks when they

[1]The subscript "A" represents the alternative hypothesis, in contrast with the null hypothesis "0".

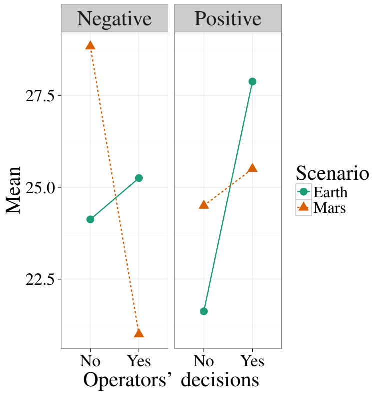

Fig. 4. Average operators' decisions in function of the Text framing and the Scenario

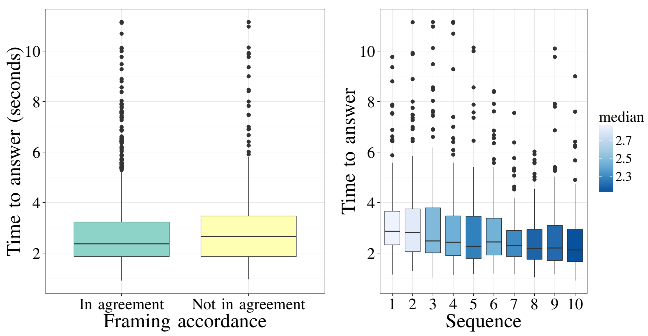

Fig. 5. Time to answer in function of the (a) text framing and the (b) sequence of missions.

were thinking about saving lives. The *positive frame*, in contrast with the negative one, was more effective in the *earthquake* scenario, confirming $H2_A$.

C. Hypothesis H3

In relation to the time to answer influence on the framing efficiency, the following hypothesis was analyzed:
- $H3_A$: The more time is devoted to answer a question the less effective is the FE.

Figure 5 (a) demonstrates that in general when the subjects answered not in accordance with the framing presented they expended more time than otherwise ($Estimate = -0.120, sd = 0.060, z = -2.004, p = 0.045$), confirming the $H3_A$. Maybe here they were using the analytical thinking mode instead of the intuitive one. Also regarding the time to answer, it is apparent from (b) a training effect from the first to the last mission executed, reducing the average time to answer, regardless of the missions, once they were randomly presented to each subject. This suggests a translation from the analytical to the intuitive thinking mode, where the FE is more effective.

D. Hypothesis H4

With the purpose of observing the satisfaction and confidence levels in function of the operators' decisions, we defined the following hypothesis:
- $H4_A$: The operators' levels of *Confidence* and *Satisfaction* will increase when they believe that they made the best choice possible in a given situation.

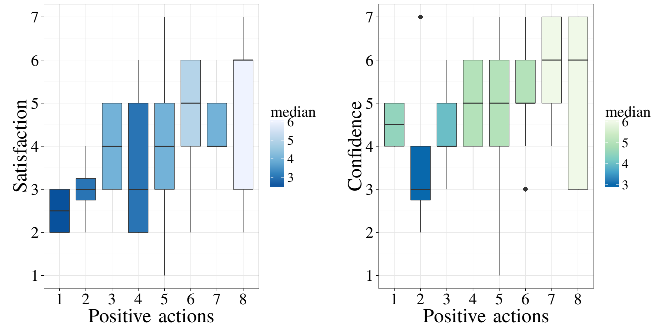

Fig. 6. Satisfaction and confidence levels in function of the operator's positive actions

Figure 6 shows these levels according to the number of positive actions executed (to release a first aid kit or to collect a rock). This criterion was used because, as already said before, it was not possible to know the results of the operators' actions (if they really helped some victims or collected some "good" rocks, respectively) at the end of a given mission. It is possible to see a significant correlation between these factors and their actions (LRT: $Satisfaction : \chi^2(2) = 163.77, p < 2.2e-16$ and $Confidence$: $\chi^2(1) = 66.064$, $p < 2.2e-16$), which indicates that we would reject $H4_0$ in favor of $H4_A$.

E. Hypothesis H5

In order to answer two research questions: (a) *What effects have Positive or Negative framed questions and Probability on the operator's decision?* and (b) *Is there a choice of Framed questions that could provide a predictable decision regardless the Probability?*, one last hypothesis was proposed:
- $H5_A$: The higher the *Probability* more the operator will be willing to say "yes" for the *Positive framing* and "no" otherwise.

Looking to Figure 7, it is apparent that $H5_A$ is corroborated (GLMM model LRT: $\chi^2(12) = 1216.1$, $p < 2.2e-16$). Also, we can see again that the positive framing is more effective in the *earthquake* scenario (*Hypothesis H2*) and the negative framing, in the *rock sampling* one. Interestingly, in the case of the *negative framing* with a "Middle-High" probability (between 51% and 63% of chance), in contrast to the *positive framing* with a "Middle-Low" probability (between 37% and 49% of chance), that are equivalent versions to present the same situation, the expected rational decision was a "NO" (not to launch a kit or collect a rock), but this only occurred in the *Mars* scenario with a negative framing. Possibly, the subjects, realizing that it was not possible to accomplish the missions with a 100% of success (releasing all eight kits or collecting eight rocks) with a threshold close to 50%, they reduced their thresholds in the next missions under the "Middle-Low" probability for the positive framing (respectively for the negative framing), as the majority of them (65.2%) reported in the questionnaire. Please, note that answer "YES" for the "Middle-Low" probability in the Positive Frame case can also be explained by the emotional commitment that influences

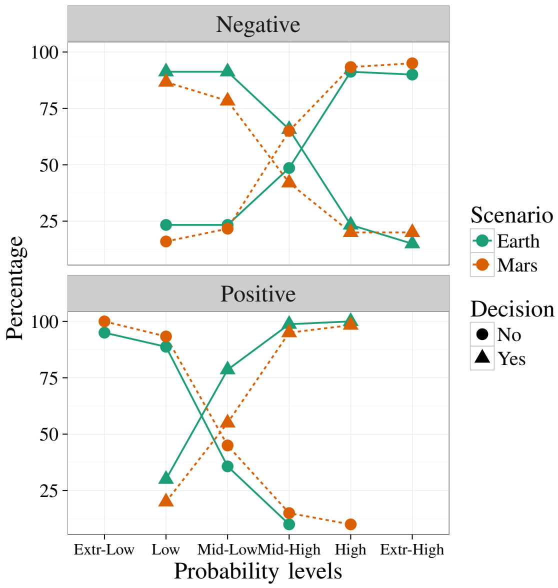

Fig. 7. Operators' decisions in function of the framing and the scenario

participants to be more risky in the *earthquake* scenario (cf. Sec. IV-B and Fig. 7).

V. Conclusion and Future Work

In order to understand what could influence human operators while making critical decisions under uncertainty and imperfect information conditions, a framing effect experiment was carried out. With a software simulator, two different scenarios were presented to the participants: (1) helping victims of an earthquake and (2) sampling rocks on Mars. During the experiment, we manipulated the framing (positive vs. negative) and the probability of the outcomes. Some hypotheses where defined to analyze the data collected. Results showed that the way the problem was presented (positively or negatively framed) and the emotional commitment (saving lives vs. collecting rocks) statistically affected the choices made by the human operators. As informed before, this is part of a work in progress, until this point 14 subjects have been observed. We are recruiting more participants and selecting new variables (*Color Visual Framing*, for instance) to observe.

The objective is, in future work, to take into account these cognitive biases to set the utility functions of the drones and to adapt dynamically human-robot interactions. The expected contributions of this work are: the introduction of a adaptive framework that, taking into account some guidelines previously defined, will be able to modify the utility functions of the drones and frame the communication with the human operator in order to optimize the mission results.

Acknowledgment

This work was supported by CNPq, Conselho Nacional de Desenvolvimento Científico e Tecnológico - Brazil.

References

[1] M. A. Hearst, "Trends & controversies: Mixed-initiative interaction." *IEEE Intelligent Systems*, vol. 14, no. 5, pp. 14–23, 1999.

[2] R. R. Murphy, S. Tadokoro, D. Nardi, A. Jacoff, P. Fiorini, H. Choset, and A. M. Erkmen, "Search and rescue robotics," in *Springer Handbook of Robotics*. Springer, 2008, pp. 1151–1173.

[3] J. Suarez and R. Murphy, "A survey of animal foraging for directed, persistent search by rescue robotics," in *Safety, Security, and Rescue Robotics (SSRR), 2011 IEEE International Symposium on*. IEEE, 2011, pp. 314–320.

[4] S. Xue, J. Zeng, and G. Zhang, "A review of autonomous robotic search," in *Electrical and Control Engineering (ICECE), 2011 International Conference on*. IEEE, 2011, pp. 3792–3795.

[5] D. Scaramuzza, M. C. Achtelik, L. Doitsidis, F. Friedrich, E. Kosmatopoulos, A. Martinelli, M. W. Achtelik, M. Chli, S. Chatzichristofis, L. Kneip *et al.*, "Vision-controlled micro flying robots: from system design to autonomous navigation and mapping in gps-denied environments," *Robotics & Automation Magazine, IEEE*, vol. 21, no. 3, pp. 26–40, 2014.

[6] F. Amigoni, N. Basilico, and N. Gatti, "Finding the optimal strategies for robotic patrolling with adversaries in topologically-represented environments," in *Robotics and Automation, 2009. ICRA'09. IEEE International Conference on*. IEEE, 2009, pp. 819–824.

[7] E. Hernández, J. d. Cerro, A. Barrientos *et al.*, "Game theory models for multi-robot patrolling of infrastructures," *International Journal of Advanced Robotic Systems*, vol. 10, no. 181, 2013.

[8] D. Portugal and R. Rocha, "A survey on multi-robot patrolling algorithms," in *Technological Innovation for Sustainability*. Springer, 2011, pp. 139–146.

[9] M. Ai-Chang, J. Bresina, L. Charest, A. Chase, J.-J. Hsu, A. Jonsson, B. Kanefsky, P. Morris, K. Rajan, J. Yglesias *et al.*, "Mapgen: mixed-initiative planning and scheduling for the mars exploration rover mission," *Intelligent Systems, IEEE*, vol. 19, no. 1, pp. 8–12, 2004.

[10] 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.

[11] 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.

[12] F. Dehais, A. Goudou, C. Lesire, and C. Tessier, "Towards an anticipatory agent to help pilots," in *AAAI 2005 Fall Symposium" From Reactive to Anticipatory Cognitive Embodied Systems", Arlington, Virginia*, 2005.

[13] 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.

[14] F. Dehais, M. Causse, and S. Tremblay, "Mitigation of conflicts with automation use of cognitive countermeasures," *Human Factors: The Journal of the Human Factors and Ergonomics Society*, vol. 53, no. 5, pp. 448–460, 2011.

[15] F. Dehais, M. Causse, F. Vachon, and S. Tremblay, "Cognitive conflict in human–automation interactions: a psychophysiological study," *Applied ergonomics*, vol. 43, no. 3, pp. 588–595, 2012.

[16] D. Kahneman, *Thinking, fast and slow*. Macmillan, 2011.

[17] G. Klein, "Developing expertise in decision making," *Thinking & Reasoning*, vol. 3, no. 4, pp. 337–352, 1997.

[18] D. Kahneman and A. Tversky, "Prospect theory: An analysis of decision under risk," *Econometrica*, vol. 47, no. 2, pp. 263–91, 1979.

[19] J. Baron, *Thinking and Deciding*. Cambridge University Press, 2007.

[20] I. P. Levin, S. L. Schneider, and G. J. Gaeth, "All frames are not created equal: A typology and critical analysis of framing effects," *Organizational behavior and human decision processes*, vol. 76, no. 2, pp. 149–188, 1998.

[21] A. Tversky and D. Kahneman, "The framing of decisions and the psychology of choice," *Science*, vol. 211, no. 4481, pp. 453–458, 1981.

[22] D. Montgomery, *Design and Analysis of Experiments, 8th Edition*. John Wiley & Sons, Incorporated, 2012.

[23] A. Agresti and M. Kateri, *Categorical data analysis*. Springer, 2011.

[24] D. Bates, M. Mächler, B. Bolker, and S. Walker, "Fitting linear mixed-effects models using lme4," *Journal of Statistical Software*, vol. 67, no. 1, pp. 1–48, 2015.

[25] S. R. Lele, J. L. Keim, and P. Solymos, *ResourceSelection: Resource Selection (Probability) Functions for Use-Availability Data*, 2016, r package version 0.2-6. [Online]. Available: https://CRAN.R-project.org/package=ResourceSelection

[26] M. Lesnoff and R. Lancelot, *aods3: analysis of overdispersed data using S3 methods*, 2013, aods3 package version 0.4-1.