Open Archive Toulouse Archive Ouverte (OATAO)

OATAO is an open access repository that collects the work of some Toulouse researchers and makes it freely available over the web where possible.

This is an author's version published in: https://oatao.univ-toulouse.fr/27716

Official URL:

To cite this version :

Hosseini, Shahrzad and Causse, Mickaël and Landgraf, Markus and Krueger, Thomas and Lizy-Destrez, Stéphanie and Dehais, Frédéric Quantifying performance in human-robotic integrated operations for spaceflight applications: a mission-driven approach. (2018) In: 69th International Astronautical Congress (IAC-18), 1 October 2018 - 5 October 2018 (Breme, Germany).

Any correspondence concerning this service should be sent to the repository administrator: tech-oatao@listes-diff.inp-toulouse.fr

Quantifying Performance in Human-Robotic Integrated Operations for Spaceflight Applications: Preliminary results

S. Hosseini[a,b,c,*], **M. Causse**[b], **M. Landgraf**[a], **T. Krueger**[a], **S. Lizy-Destrez**[b], **F. Dehais**[b]

[a] *European Space Research and Technology Centre, European Space Agency, Keplerlaan 1, 2201 AZ, Noordwijk, the Netherlands*, {Shahrzad.Hosseini, Markus.Landgraf, Thomas.Krueger}@esa.int

[b] *Institut Supérieur de l'Aéronautique et de l'Espace, 10 Avenue Édouard-Belin BP 54032 - 31055 Toulouse CEDEX 4, Toulouse, France*, {Frederic.Dehais, Stephanie.Lizy-Destrez}@isae-supaero.fr

[c] *Airbus Defence and Space, Sebaldsbrücker Heerstraße 235, 28309, Bremen, Germany*

[*] *Corresponding Author*

As the global space exploration community moves towards the exploration of the Moon and beyond, so must the preparation of crew operations. Given the distinct operational requirements of human lunar exploration scenarios compared to the on-going ISS operations, the study of human-robotic integrated operations (HRIO) is key in the endeavour of mankind's return to the Moon and enabling a sustainable exploration strategy. Moving beyond qualitative performance assessments of HRIO, one particular issue for exploration destinations are time-delay conditions, and therefore the step must be made towards quantifying performance in these operations such that both crew and system designs are prepared accordingly. Based on the pilot study previously presented in Hosseini et al. [1], the current paper presents a mission-driven experiment campaign set up to study human performance regarding the ESA-led HERACLES mission, a proposed sub-scale tele-operated demonstrator mission aiming to prepare international partners for human lunar missions. A targeted geological site in a lunar crater is set up in an analogue environment at ESA allowing a rover to be tele-operated. The knowledge gap in HRIO is challenged in this study, since reaching this level of human-robotic partnership requires an unprecedented understanding of the interaction between the human and robotic system. The approach to fill this gap is to quantify objectively the HRIO performance for spaceflight applications, by studying human and robotic elements as two separate yet cooperating systems. 16 participants were instructed to drive the rover through an obstacle course using a controller and camera as interface to the rover. Three mission-driven time-delay conditions were applied to simulate different control configurations, i.e. 3.5s, 0.5s, and 0s representing control from ground, cis-lunar space, and lunar surface, respectively, assuming the rover is driving on the lunar surface. The experiment is set up such that human performance metrics are acquired following a neuroergonomics approach, focusing on the cardiovascular activity to infer participant's mental workload, and ocular behaviour, to measure attentional abilities. In parallel, robotic metrics are acquired through the hardware and software output of the rover. Studying human and robotic data output recorded in parallel allows quantification of the level of mental workload under the delay conditions and the resulting effects on the HRIO performance. This approach is believed to advance the level of detail and understanding of HRIO as known to date, subsequently identifying the key elements to prepare astronauts for future missions.

Keywords: Space Exploration, Human-Robotic Integrated Operations, Lunar Operations, Analogue Test, Performance, Neuroscience, Neuroergonomics, Human Factors

1. Introduction

Agencies are preparing the future of spaceflight by planning future missions to the lunar vicinity and to the lunar orbit. An ESA-led robotic demonstrator mission called HERACLES [2] aims to send humans and rovers to the Moon in order to extend the reach of mankind into our Solar system, and to gain unprecedented knowledge by doing so.

Human-Robotic Integrated Operations is believed to be key for these missions. A study by Crawford in 2012 [3] argued that the partnership of both human and robotic system go farther than either one could do individually, and this study aims to quantify this argument in terms of performance data. The

Copyright © 2018 by the International Astronautical Federation (IAF). All rights reserved.

List of Acronyms

ECG	Electrocardiogram
ESA	European Space Agency
ESTEC	European Space Research and Technology Centre
HERACLES	Human-Enabled Robotic Architecture and Capabilities for Lunar Exploration and Science
HMI	Human-Machine Interface
HRIO	Human-Robotic Integrated Operations
ISECG	International Space Exploration Coordination Group
ISS	International Space Station
LOP-G	Lunar Orbiter Platform - Gateway
TLX	Task Load Index

focus of the study is on both robotic- and human data, analysing both the engineering software and the "human software" through an neuroergonomics approach.

The goal of this study is to simulate a scenario in which an operator, this can be either an astronaut in space or flight controller from a mission control centre on Earth, is given control of a rover which must drive on the Lunar surface in order to observe the site prior to its sampling operations. This study prepared an analogue site for this in the ERASMUS Highbay at ESA ESTEC, in order to answer the following questions.

- Does increasing time-delay result in an increase of perceived pressure, i.e. mental-, physical-, and temporal demand, by the operator?
- Does increasing perceived pressure lead to degradation in human performance in human-robotic integrated operations?

An experiment was conducted in the analogue site, with pilot trainees from the Royal Netherlands Air Force. The preliminary results of this experiment are presented in the scope of this paper.

2. Material and methods

2.1 *Subjects and task instructions*

Sixteen students of the Royal Netherlands Air Force in the selection process to becoming military pilots were invited for a tele-operated driving experiment at ESA ESTEC. The participants were in the range of 18-28 year old, and although the invitation was open to both men and female trainees, the applicants for this experiment were all male. The participants were welcomed into a dedicated room in the ERASMUS Highbay area of ESTEC and asked to take place in their workstation consisting of a race chair with a large screen in front at eye-level. A FAROS electrocardiogram (ECG) was applied on their chest with three electrodes to record the heart rate, and a Tobii eye-tracker was applied on their eyes to track their eye-movements.

The rover was positioned in the Highbay at just meters distance from the participants, but the rover was not in their field of sight, neither prior to the experiment nor during the experiment. The area in which the rover was operating was closed off such that the participants did not get a view of the site. This is done to simulate the scenario in which future operators of lunar rovers will also not have seen the site when operating the rover. For visibility of the task that was to be performed, a camera was mounted on the rover that was in a fixed position and gave a wide-angle view of the forward direction of the rover's path.

The instruction given to the participants was to drive the rover from its starting position, through an obstacle course, avoiding the numerous mock rocks, to its final destination (also the starting position) which was indicated by a rectangle box. In order to find their way through the course, they were asked to follow the white arrows that were on the ground. In order to achieve an optimal data collection, the participants were asked to sit calmly and limit their speaking to a minimum. The same course was driven by each participant for a total of three times, under different time-delay conditions which are identified in the *Independent Measures*.

2.2 *Apparatus*

The participants, acting as operator in the experiment, were provided with an HMI with which they could control the rover from a distance.

Rover and HMI

The rover that the operators operated is the Interact Centaur Rover, an in-house developed experimental rover by the ESA's Human Robotic Interaction Lab [4]. The rover consists of a mobile platform (all four wheels steered and powered), with two robot arms that were kept inactive for this experiment, and a stereo camera that was locked in its position. The HMI is a simple hand controller that is often used for gaming, with one button active which allows the operator to control the rover to the front, back, left, and right.

Copyright © 2018 by the International Astronautical Federation (IAF). All rights reserved.

2.3 *Independent measures*

As this study aims to analyse the effect of time-delay on the operator's performance, the set of conditions of this study consists of three different time-delay scenarios for controlling a rover that is assumed to be driving on the Lunar surface.

- Condition 1: 0s delay, assumes the operator to be on the Lunar surface either near or in the rover
- Condition 2: 0.5s delay, assumes the operator to be in cis-lunar space, i.e. in a station orbit in the Lunar vicinity
- Condition 3: 3.5s delay, assumes the operator to be in a mission control room, on Earth

2.4 *Dependent measures*

The dependent measures are listed in Table 1.

Table 1: Dependent measures

Measure	**Symbol**	**Description**
Safety	$n_{collisions}$	Number of collisions [-]
Performance	TOC	Time of completion [s]
Subjective	TLX	Workload assessment survey
Subjective	SA	Mapping exercise
Stress	ECG	Electrocardiogram data
Focus	*Tobii*	Eye-tracker data

2.5 *Procedure*

Prior to the start of the experiment, each participant was briefed about the task they were about to perform, and some practical information regarding the sensors was given. They were informed that they must control the rover using the manual controller, and to follow the white arrows from starting position, through the course, to the end position. Moreover, the participant were informed multiple times during the experiment, the experimenter may give them a code that they are asked to repeat. Another instruction was to be efficient with the time, because of the battery duration.

The exercise started with a training period of 10 minutes. This duration of this training time was concluded in the pilot study [1] which was held in preparation for this experiment, based on the time performance results that showed a converging behavior after approximately 10 minutes of familiarisation with the system. The training period consisted of controlling the rover under the three delay conditions that the experiment cover, in order to prepare the participants for each condition. Furthermore, a parking exercise was given in order to practice the positioning of the rover. It is worthwhile noting that the training are was different than the main course, to avoid memorisation of the course which may influence the data output.

After the training period, the rover was driven to its starting position by the experimenter and once at its starting position, the control was handed over to the participant. After the start signal was given by the experimenter, the participant started to drive the rover through the course while avoiding the obstacles it was faced with. For a total of three times, at three fixed locations on the course, the experimenter read a 7-digit code and asked the participant to repeat the code in the same order in which it was presented to them.

After each round, the participant was presented a digital NASA TLX questionnaire which they were asked to fill in, such that the workload could be measures for each time-delay condition. Upon completion of the questionnaire the next round would start, until all three rounds had been completed. After the last experiment round, the participant was provided with a map which showed a schematic sketch of the rover and the obstacles, plus false positive obstacles. The participant was asked to identify by memory the obstacles that it had faced during the experiment, and the route that the rover had driven.

As stated, the course was driven three times, under three different conditions each simulating a control condition. The order of conditions for the participants are counter-balanced. The conditions and their reasoning are presented next.

Condition 1

Condition 1 simulates the situation in which the operator is on the Lunar surface, controlling a rover that is on the Lunar surface as well. This can be either a smaller rover with direct control, or a pressurised rover which seats the operator. Because of the direct control that is assumed, the time-delay is 0s.

Condition 2

Condition 2 simulates the situation in which the operator is in a station (from now referred to as LOP-G) in cis-lunar space, orbiting in the vicinity of the Moon. Meanwhile, the rover is also on the Lunar

Copyright © 2018 by the International Astronautical Federation (IAF). All rights reserved.

surface, with a communication link to the LOP-G. At its highest communication link, considering the orbit in a Near-Rectilinear Halo Orbit and approximately 60.000km from the Lunar surface, the time-delay the operator is faced with is 0.5s. Depending on the communication link and the position of the LOP-G with respect to the rover's position, this time-delay can in the order of minutes. In the scope of this study, the best-case scenario is considered.

Condition 3
Condition 3 is the simulation of the situation in which the operator is in a mission control room on Earth, controlling the rover that is on the Lunar surface. The operator has to face the highest time-delay of the three conditions, i.e. 3.5s at its highest communications point. The delay of this condition can also be in the order of minutes but as explained earlier, the best-case scenario is considered in the scope of this study.

2.5.1 *Performance observation and measurement*

In addition to the experiment who was in the control room with the participant at all times, a support engineer accompanied the rover in the Highbay. This was done in order to have full-time observation of both the human operator and the rover. The latter is done as a safety requirement to avoid any collisions with any other hardware in the highbay, and therefore the support engineer had a control which could overrule the participant's contoller, in case the rover had to make an emergency stop.

3. Theory and Results

Data is obtained from both technically performed and perceived output, but in this paper the preliminary results of this study are presented. The sensor data, measured by the ECG and Eye-tracker, will be presented in a more elaborate version of the analysis and for a larger subject pool. The data analysis is presented for the time performance, the workload, and the situational awareness.

3.1 *Time performance analysis*

In the analysis of the time performance, the time of completion of the experiments rounds are studied. Each participant's time of completion is tracked, for each of the time-delay conditions in which it is performed. Based on the results of the sixteen participants, the mean value for C1, C2 and C3 are presented in figure 1. It can be seen that for an increased time-delay, the mean time of completion increases.

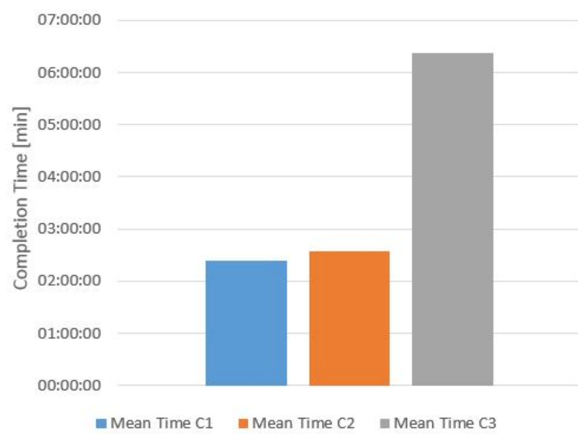

Fig. 1: Time of Completion

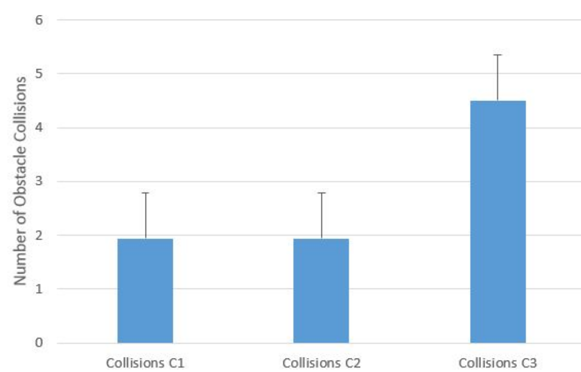

Fig. 2: Obstacle Collisions

This result may be supporting the hypotheses, but that cannot be determined yet without the sensor data. Though the time of completion increases with time-delay, this does not yet proof degradation of human performance in the interaction between the operator and the rover. A further analysis, with a greater sample size and full analysis of the ECG and Eye-tracker data is expected to provide a complete answer to the hypotheses.

Furthermore, the number of obstacle collisions per condition are analysed, and presented in figure 2. It can be seen that for C1 and C2, the mean number of collisions is relatively low, and equal amongst the two conditions. For C3 however, a significantly higher number of obstacle collisions is presented, about two times the mean value of C1 and C2. This was as expected, since C3 required a much higher demand in route planning and anticipation of rover movements.

Copyright ©2018 by the International Astronautical Federation (IAF). All rights reserved.

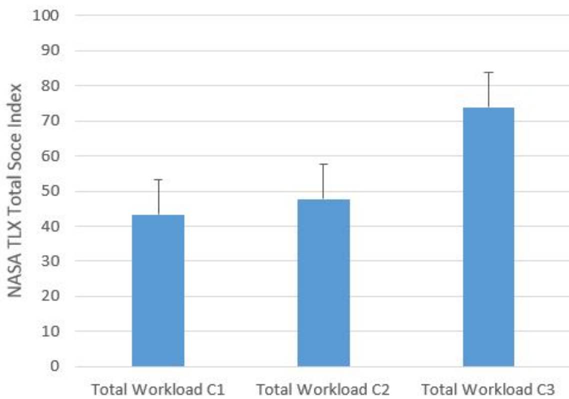

Fig. 3: Total Workload

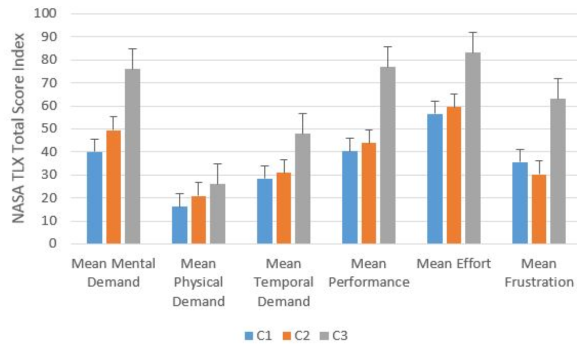

Fig. 4: Workload

3.2 *Workload analysis*

The workload of the operator in the human-robotic interaction is determined using the NASA Task Load Index (TLX), which is a method to quantify the perceived workload of the operator. In this questionnaire, the operator is asked to rate their perception of their performed task after it has been completed. This is done in the category of *mental demand*, *physical demand*, *temporal demand*, *performance*, *effort* and *frustration*. The questionnaire consists of two parts, the first asking the participant to rate the aforementioned six categories on a scale of 0 - 100. The second part aims to identify the weight of each category, and therefore asks the weight of each category with respect to another. As a result, for each category a total given score and its weight are presented as output, and moreover the total workload is presented.

The results of the workload total workload are presented in figure 3, and as can be seen, the perceived workload increases as the time-delay increases. As expected, the differences between C1 and C2 are not significantly large, since the two conditions differ 0.5s in value. The total workload of C3 on the other hand does show a significant difference with C1 and C2. This was as expected, since a significant time-delay of 3.5s results in a more intensive situation for the operator.

Looking at the six categories separately in figure 4, it can be seen that for the mental demand, the perceived score increases with the time-delay. This is as expected, due to the extra mental planning that is required when the operator must plan ahead for a number of seconds, as compared to an instantaneous response from the rover. The difference between C3 and the other two conditions is significant, as also the value of the time-delay is significantly greater.

Although the physical demand is rated fairly low in general, it also shows a slight increase in the perceived workload. But compared to the temporal demand, which aims to assess the perceived time pressure by the operator, the latter shows a greater significance for C3 in comparison with C1 and C2.

Interesting is to see the fact that as the time-delay, and thus from the results also the mental demand increased, the results also show an increased score of the perceived performance. The operators clearly indicate a significant increase of increased effort with increased time-delay, and with that also comes an increased level of frustration. But even with the latter being taken into account, the perceived performance, i.e. the perception of the operator of how well they did their tasks, also increased as more demand was required.

3.3 *Situational awareness analysis*

For the assessment of the situational awareness of the participants, a map was presented to the participants with a schematic sketch of the rover, the obstacles, and a number of false positive obstacles, as explained earlier in the experiment procedure. The participants were presented the map at the end of the experiment, i.e. upon completion of all three rounds, and were asked to indicate 1) the path they had driven, and 2) the obstacles they encountered. All participants were presented the same map, and an analysis of correct answers, false negatives and false positives was performed.

The course that was set up for the experiment was divided in two parts, a fact unknown to the participants, and the difference between the two parts were the levels of obstacle complexity. The first part of the course consisted of two-fold obstacles, i.e. two obstacles that were located on a horizontal line with respect of the approaching rover. This first part of the course is called Area 1. The second part consisted

Copyright ©2018 by the International Astronautical Federation (IAF). All rights reserved.

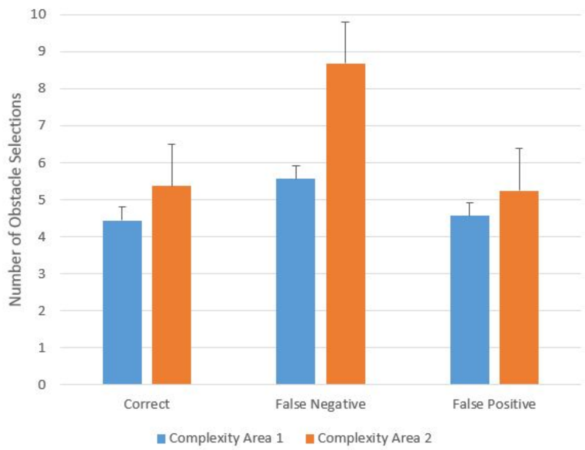

Fig. 5: Situational Awareness

of three-fold obstacles which were three mock rocks located in a triangle such that the rover had to make more maneuvers to get through it. This part of the course is called Area 2. An analysis is made on the difference in performance in Area 1 and Area 2.

From the results in figure 5 it can be seen that first of all, a higher number of correct answers were given for Area 2, which indicates that a more complex area results in a higher situational awareness of the operator. It must be considered that in Area 2 a higher number of rocks were placed, forming the three-fold obstacles, and therefore the chance of identifying a correct rock is also higher compared to Area 1 in which less rocks are present. The most significant difference can be seen in the false negative answers, and it can be seen that in a more complex area, also more rocks are overlooked. Studying this fact and the previous one in parallel, and simultaneously looking at the difference in significance of both, it can be concluded that indeed the number of rocks present in Area 1 may have an effect on the amount of correct answers. In fact, the conclusion is drawn that in a high-complexity terrain as expected on the lunar surface, the performance in situational awareness may be lower because there is more information to absorb. Regarding the false positives, in Area 2 a slightly higher number of rocks are added to the map beyond the ones that were actually present.

4. Conclusion and recommendations

With the aim to gather data on human performance in human-robotic integrated operations, multiple conclusions are drawn from the experiment that has been conducted. As seen from the data, multiple methods of performance assessment has been applied, and the conclusions are as follows. First of all, in the time performance data it was shown that for an increased time-delay, the mean time of completion increases. Furthermore, in the NASA TLX data an increased perceived workload was assessed as the time-delay increased. A point to note was the fact that in addition to the demand, the perceived performance also increased as more demand was required.

The number of collisions proved to be higher for a greater time-delay condition. The results showed that two times more collisions were made for the high-delay condition, as compared to the two low-delay conditions.

From the situational awareness assessment it was seen that in a high-complexity terrain as expected on the lunar surface, the performance in situational awareness may be lower because there is more information to absorb. It must be noted that this experiment is designed such that the human performance can be measured without any aid that improves or affects the performance. In spaceflight missions there will be sensors and other aids which will support the operator, especially in high-complexity areas. These may be extra cameras, side cameras, rear cameras, collision avoid systems etc.

From the results that were presented it can be stated that the conclusions do support the hypotheses, but they do not prove it yet. This is because the major part of the data analysis is not taken into account in the scope of the representation of the preliminary results, and moreover the full study covers a total of 40 participants and therefore more data will be available for analysis.

Acknowledgements

First and foremost, the Royal Netherlands Air Force is thanked greatly for their support and contribution in terms of manpower to this study. The effort of Colonel de Jongh of coordinating the training teams withing the Air Force led to a successful realisation of the experiment, and Captain Willems' recommendations made a great contribution to this academic cooperation. Thanks also go to the support engineers Max Braun and Pepijn Versloot for monitoring the rover during the experiments. And lastly, thanks go to the Human Robotic Interaction Lab of ESA ESTEC for the provision of the rover and the in-house support.

References

[1] S. Hosseini et al., Pioneering steps towards future human-robotic operations performance, Confer-

Copyright © 2018 by the International Astronautical Federation (IAF). All rights reserved.

ence Paper at the International Astronautical Congress (2017)

[2] M. Landgraf et al., HERACLES Concurrent Design Facility Study internal Report, (2015), Noordwijk, ESA CDF.

[3] I. Crawford, Dispelling the myth of robotic efficiency: why human space exploration will tell us more about the Solar System than will robotic exploration alone, Astronomy and Geophysics (2012), Volume 53, pp. 2.22-2.26.

[4] ESA Telerobotics and Haptics Laboratory, `http://esa-telerobotics.net/` [Online]