Towards a Measure of Situation Awareness for Space Mission Schedulers

Tamsyn Edwards[1], Summer L. Brandt[2], Jessica J. Marquez[2]

[1] San Jose State University at NASA Ames Research Center, Moffett Field, CA 94035
[2] NASA Ames Research Center, Moffett Field, CA, 94035

{Tamsyn.E.Edwards, Summer.L.Brandt, Jessica.J.Marquez}@nasa.gov

Abstract. The success of space flight missions relies on support provided to astronauts through specialist knowledge of ground-based personnel. One of the many essential tasks that ground personnel provide is the scheduling of flight crew daily activities. Future long duration exploration missions will require astronauts to assume planning and scheduling responsibilities in order to facilitate increased autonomy from ground support. Although situation awareness is critical to the scheduling task, a sufficient measure for this domain has not been developed. This paper documents the approach and process by which the authors developed a framework for measuring situation awareness in space mission schedulers and presents the measure applications' initial results.

Keywords: Human factors in spaceflight · Self scheduling · Situation awareness measurement · Human performance

1 Introduction

The success of space flight missions relies on the specialist knowledge of a large team of ground-based personnel who are responsible for essential tasks to support astronauts. One such task is scheduling daily activities for astronauts. Mission schedulers create schedules for the duration of the mission, ensuring that activities are scheduled efficiently and effectively. The scheduling role is therefore essential to mission success. As NASA moves toward long-duration exploration missions (e.g. Mars), pre-scheduling all astronaut activities by ground personnel becomes inefficient. Missions that operate further from Earth will have high communication latencies between the spacecraft and Mission Control Center. Instead of waiting for guidance from ground personnel, astronauts should be able to manage their own schedule, enabling them to use their time more efficiently and effectively. This requires astronauts to assume partial responsibilities of scheduling, allowing crew to be more autonomous from ground support [1]. Unfortunately, astronauts will not have the same specialist knowledge ground personnel have to develop effective schedules.

Previous studies have identified situation awareness (SA) as one critical factor for developing effective schedules [2]. However, a gap exists in research related to understanding and defining SA within the scheduling context, and development of a suffi-

cient measure for this domain. This paper documents the approach and process by which the authors developed a framework for measuring SA in space mission schedulers and presents initial results on the sensitivity of the developed SA measure to detect time to answer accurately.

2 Situation Awareness Overview

SA is typically defined as "the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future" [3, p97]. Inherent in [3]'s commonly cited definition is the assumption that SA is categorized into three distinct levels: Level 1 is a situation assessment of the environment, Level 2 involves understanding of the current surroundings/situation and Level 3 includes the prediction of an upcoming situation or surroundings [4]. Many domains have utilized this definition, often tailored to the specific profession and environment under consideration [5]. However, there is a gap in the scheduling literature in both definition and measurement of SA. In order to adequately measure SA, it is first necessary to adequately define SA in relation to the scheduling role.

2.1 Understanding SA in Scheduling

Mission schedulers are responsible for creating astronauts' schedules for the duration of the mission. They need to ensure that activities are scheduled according to requirements, such as a specific time of day or priority of an activity, while preventing scheduling conflicts [6]. Significant experience is needed to create an effective and efficient schedule as ground-based personnel must be aware of elements that will affect the schedule, such as the spatial location of activities onboard the ISS and crew knowledge required for specific assignments. As such, creating a mission schedule for ISS generally takes several weeks to months.

To gain further insight into the meaning of SA specific to schedulers, interviews were conducted with seven subject matter experts who were current or recently retired space mission schedulers. Semi-structured interview questions focused on several topics, including scheduling task steps, strategies, and information required for an effective schedule. After transcription and application of thematic analysis, patterns emerged in the data that related to the three levels of SA: (i) Level 1 – 'Perception' (knowledge): of factors pertinent to the scheduling problem - e.g., activity requirements, activity priority, number of activities; (ii) Level 2 - 'Comprehension': an understanding of those constraints and what it means for the scheduler – e.g., it is most important to plan high priority activities (iii) Level 3 - 'Prediction': How to plan to avoid inefficiencies - e.g., schedule high priority activities first. Interestingly, schedulers reported that experience was needed in order to build awareness and understanding of all the required elements to build an effective schedule (mapping to Levels 1 and 2 SA). A specific example included 'formal' and 'informal' constraints. Formal constraints were factors that were documented officially and explicitly for schedulers to follow (Table 1). Informal constraints were described as factors that were essential to an effective schedule but were not documented or trained. Instead, development of

awareness for these constraints arose from experience (Table 1). Findings were not only relevant to understanding the development of SA in mission scheduling, but also highlighted the specific areas that may be objectively measured with an SA measurement technique.

Table 1. Formal and informal scheduling constraints considered by schedulers (Level 1 SA)

Formal scheduling constraints		Informal scheduling constraints
Activity priority		Spatial constraints onboard ISS
Activity duration		Spatial relationships of activity locations
Maximizing crew time		Activity knowledge
Activity constraints:		Flexibility of activity constraints
a.	Ordering Constraint	Equipment needed (impacts duration)
b.	Time Range Constraint	Crew preferences/personality
c.	Requires Constraint	Crew qualifications
d.	Overlapping Activities	Set up time (impacts duration)
		Crew comfort

2.2 Scheduling-specific SA Measure Development

Several SA measurement methods are documented in the literature, spanning direct objective and subjective measures, and indirect or implicit measures (for a full review please see [7]). In relation to mission scheduling, the Situation Present Assessment Method (SPAM) [8] appears to offer several advantages, including that the measure can be used without creating interruptions, and results have been found to be less influenced by memory confounds [8]. However, other aspects of SPAM are not particularly suited to the specifics of the scheduling task; the scheduling activity is less dynamic than many other domains, including the air traffic control domain in which SPAM was originally developed. For the current study, elements of SPAM were utilized to develop a probe-based, direct measure of SA, specifically relevant for schedulers. Binary response options were provided. Utilizing the findings from the mission scheduler interviews, probe questions were developed to assess formal scheduling constraints. Questions focused on the current situation (Levels 1 and 2 SA; Did you leave 3 or more activities unscheduled?) as well as future prediction (Level 3 SA; If COMM was suddenly unavailable before 10:00, would any activities be in violation?). Our measure departed from SPAM in two manners: probing questions were given at the end of a scheduling problem task (to minimize interruptions) and workload measure of SPAM was not necessary. The experiment described below featured the initial testing of the measure.

3 Measure Implementation: Experimental Methodology

For the current study, two independent variables were used: Type of Constraint on Activities (see below for a full description of constraints) and Percentage of Constrained Activities (33% or 66% of activities), resulting in a 4 x 2 within-subjects design. A baseline trial was run as the first trial for all participants and contained no

manipulations; a latin square determined the order for the remaining eight trials. Following each scheduling task trial (similar to [2]), participants completed three questions pertaining to SA. Participants were instructed to answer the questions as quickly and accurately as they could and were permitted to refer back to the completed schedule. Twelve participants (4 Female, 8 Male) naive to the scheduling task and scheduling platform (Playbook timeline tool) volunteered to participate in the study. Participant age ranged between 35 to 64 years, and all participants held a bachelor's degree or higher.

The web-based Playbook timeline tool functioned as the study platform (see [1, 9] for an overview). Participants were given a set of activities each designated as low, medium or high priority and asked to schedule these activities among fourthree crew members. Participants were asked to schedule according to activity priority. Some activities included rules, known as constraints. The following constraints were included in the study:

- *Time Range Constraint (T)* limits the time of day an activity can be scheduled (e.g., Activity A must start no earlier than 0900 and end no later than 1030);
- *Requires Constraint (R)* states that the activity needs to have a particular resource available (e.g., Activity requires communication availability);
- *Claim Constraint (C)* describes a specific piece of equipment required for a particular activity (e.g., Activities A and B both claim a treadmill, therefore cannot be scheduled at the same time).
- *Ordering Constraint (O)* describes when an activity should be scheduled in relation to another activity (e.g., Activity A must be scheduled before Activity B);

4 Results

Descriptive statistics were calculated for the percentage of correct responses to SA questions. Overall, participants answered 71.60% of all questions correctly. Statistical analyses resulted in no significant differences between conditions. Descriptive statistics revealed a data trend suggesting that for 33% constrained activities, T33 (M=69.44%) and R33 (M=66.67%) were on average answered less accurately than C33 (M=86.11%) and O33 (M=72.22%). Percentage of responses for future-focused questions (M=70.83%) were less accurate on average than present-focused questions (M=74.17%), although the difference was not significant. Additional analyses examined time in seconds for accurate responses to SA questions. Where a participant did not answer any SA questions correctly for a specific condition, different sample sizes were recorded. A review of descriptive statistics identified that overall, response latencies were longer for the Claim and Ordering constraint conditions compared to Time Range and Requires Constraint condition. (Table 2).

Friedman's ANOVA confirmed that significant differences in accurate response time to SA questions existed between conditions ($\chi^2(8)=20.68$, $p<0.01$). Post hoc analyses revealed that accurate response times were faster in T33 ($z=-2.89$, $p<0.05$) and R66 ($z=2.56$, $p<0.05$) than the C33 condition. Accurate responses were also significantly faster for T33 ($z=-3.11$, $p<0.05$) and R66 ($z=2.78$, $p<0.05$) compared to

O33. Finally, response times were significantly faster when compared to O66 for B ($z=-3.56$, $p<0.01$), T33 ($z=-4.33$, $p=0.001$), R33 ($z=-3.67$, $p=0.005$), R66 ($z=-4.00$, $p<0.005$), and D66 ($z=-2.67$, $p<0.05$). Results suggest a significant main effect of constraint type on SA response latency, although not of percentage of constraint. A Wilcoxon signed rank analysis also revealed that accurate probe responses were significantly faster for present-focused SA probes ($M=12.88$, $SD=6.38$) the future-focused SA questions ($M=20.42$, $SD=11.73$), ($z=2.51$, $p<0.05$).

Table 2. Mean duration to respond accurately to SA questions (in seconds) by condition

Constraint Condition	n	Mean	SD
Baseline (B)	11	10.06	6.32
Time Range - 33% (T33)	12	9.89	6.20
Time Range - 66% (T66)	12	9.77	4.80
Requires - 33% (R33)	12	9.86	5.98
Requires - 66% (R66)	11	7.56	2.29
Claim - 33% (C33)	12	13.59	6.18
Claim - 66% (C66)	12	15.58	14.82
Ordering - 33% (O33)	12	21.59	24.63
Ordering - 66% (O66)	11	21.77	14.62

3. Discussion and Conclusions

A measure of SA was developed for space flight mission scheduling based on qualitative data from expert mission schedulers. The measure was utilized in an experimental study that replicated a scheduling task. A within-subjects design study investigated the effect of type of activity constraint and percentage of constrained activities on SA, as inferred from the measure. No significant effects were found for type of constraint, or percentage of constraints, on response accuracy. This result was unexpected as more constrained activities were predicted to increase complexity, and therefore, negatively affect response accuracy. One explanation for this finding is that as the participants could see their completed schedule, the likelihood of accurate answers was increased. In addition, the small sample size could have created a Type II error. Finally, the percentage of constraints condition may not have increased complexity enough to affect SA.

Results did reveal significant differences for latencies of accurate responses to SA questions, indicating that there was a main effect of constraint type on SA. Differences in response time could be an indicator of differing complexities of scheduling activities with specific types of constraints; for example, some constraints may be easier to understand than others. In line with expectations, significant differences in response times were found between present-focused and future-focused questions, suggesting that participants naive to scheduling did not develop Level 3 SA (prediction). This finding has important implications for facilitating the development of scheduling-based SA in astronauts as we assume limited scheduling experience prior to flight.

In relation to measure development, despite the adaptation of the SPAM methodology, the measure was successfully applied to a novel task and naive schedulers, for a task that is not dynamic. Initial results therefore seem to support the potential utility of this measure. A limitation of this study was that, for mission schedulers, informal constraints directly affect how to plan to avoid scheduling inefficiencies. The current study operationalized this for the naive participants and provided them with the strategy to schedule high priority activities first; for expert schedulers the task would be much more complex.

Future research is needed to investigate the ecological validity of the findings in relation to expert mission schedulers. For the wider domain, it is essential that future research investigate the process by which mission schedulers build SA from informal constraints and develop supportive aids to facilitate non-expert schedulers (such as astronauts) in the task.

Acknowledgements

This research was funded by the NASA Human Research Program's Human Factors and Behavior Performance Element (NASA Program Announcement number 80JSC017N0001-BPBA) Human Capabilities Assessment for Autonomous Missions (HCAAM) Virtual NASA Specialized Center of Research (VNSCOR) effort. The authors would like to thank Candice Lee, Megan Shyr, Casey Miller and John Karasinski for their assistance in collecting this research data.

References

1. Marquez, J.J., Hillenius, S., Deliz, I., Kanefsky, B., Zheng, J., Reagan, M.: Increasing Crew Autonomy for Long Duration Exploration Missions: Self-Scheduling. In: 2017 IEEE Aerospace Conference, pp. 1--10. IEEE (2017)
2. Lee, C.N., Marquez, J.J., Edwards, T.E.: Crew Autonomy through Self-Scheduling: Scheduling Performance Pilot Study. In: AIAA SciTech Forum, pp.1578 (2021)
3. Endsley, M.R.: Design and Evaluation for Situation Awareness Enhancement. In: Proceedings of the Human Factors Society Annual Meeting, 32(2), pp. 97--101 (1988)
4. Endsley, M.R.: Toward a Theory of Situation Awareness in Dynamic Systems. Hum. Factors 37, 32--64 (1995)
5. Fioratou, E., Flin, R., Glavin, R., Patey, R.: Beyond Monitoring: Distributed Situation Awareness in Anesthesia. Brit. J. Anaesth. 105(1), 83--90 (2010)
6. Barreiro, J, Jones, G., Schaffer, S: Peer-to-peer Planning for Space Mission Control. In: 2009 IEEE Aerospace Conference Proceedings. pp.1--9. IEEE (2009).
7. Endsley, M.R.: A Systematic Review and Meta-analysis of Direct Objective Measures of Situation Awareness: A Comparison of SAGAT and SPAM. Hum. Factors 63(1), 124--150 (2021)
8. Durso, F.T., Truitt, T.R., Hackworth, C., Crutchfield, J., Nikolic, D., Moertl, P., et al.: Expertise and Chess: A Pilot Study Comparing Situation Awareness Methodologies. In Garland, D.J., Ednsley, M.R. (eds.) Experimental Analysis and Measurement of Situation Awareness, pp. 295--304. Embry-Riddle Aeronautical University Press (1995)
9. Marquez, J.J., Pyrzak, G., Hashemi, S., Ahmed, S., McMillin, K., Medwid, J., Chen, D., Hurtle, E. Supporting Real-Time Operations and Execution through Timeline and Scheduling Aids. In: 43rd International Conference of Environmental Systems, pp.3519. ICES (2013)