What is the role of Usability and Trust in Autonomy?

Jessica J. Marquez, Ph.D.
NASA Ames Research Center

jessica.j.marquez@nasa.gov

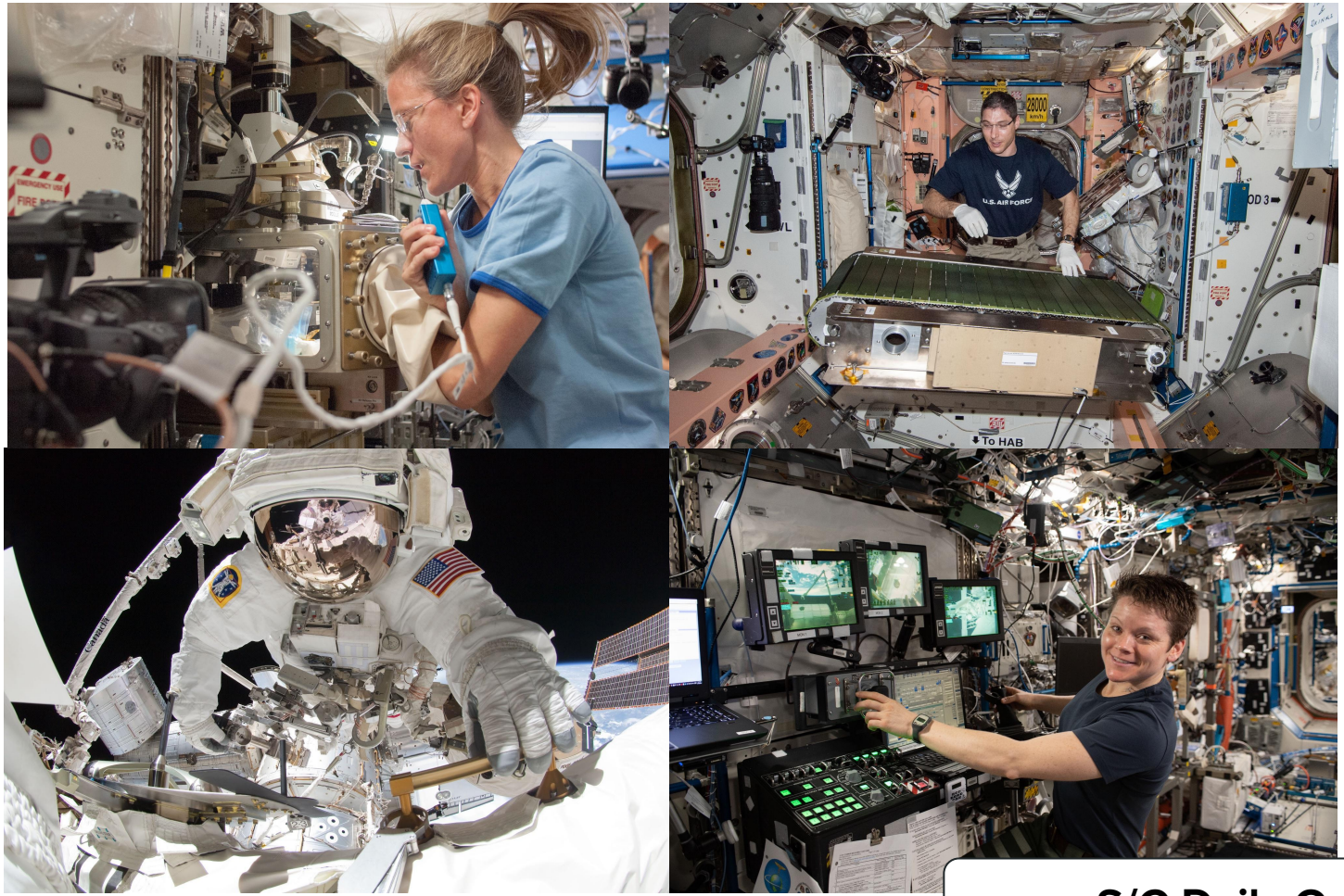
Position: Trust in autonomy is enabled by systems with high usability.

NASA Playbook Case Study

State of the Art of Human Mission Operations

International Space Station (ISS)

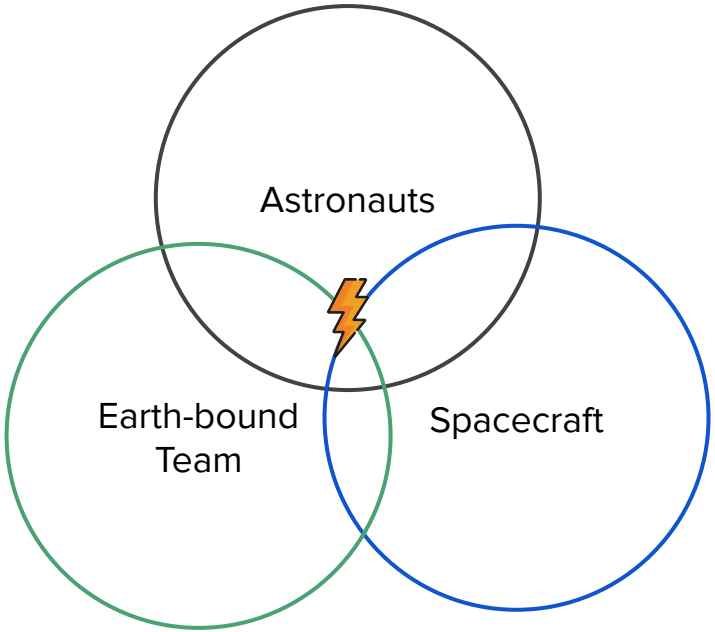
Mission Control Center

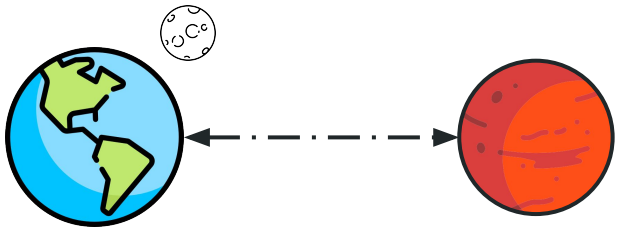


S/G Daily Ops


Human spaceflight operations needs to evolve.

Towards Earth Independent Operations





Communication transmission delays, limited bandwidth, and periods of no communication require NASA to change operations towards one that is more Earth-independent.


More automation, robotics, & autonomous
systems to enable crew autonomy.

Enabling Crew Autonomy through Self-Scheduling

A software tool that enables the crew to manage their own schedule.

- Crew contributes their insight on how to schedule & reschedule activities.
- Minimizes idle time waiting for Mission Control responses.

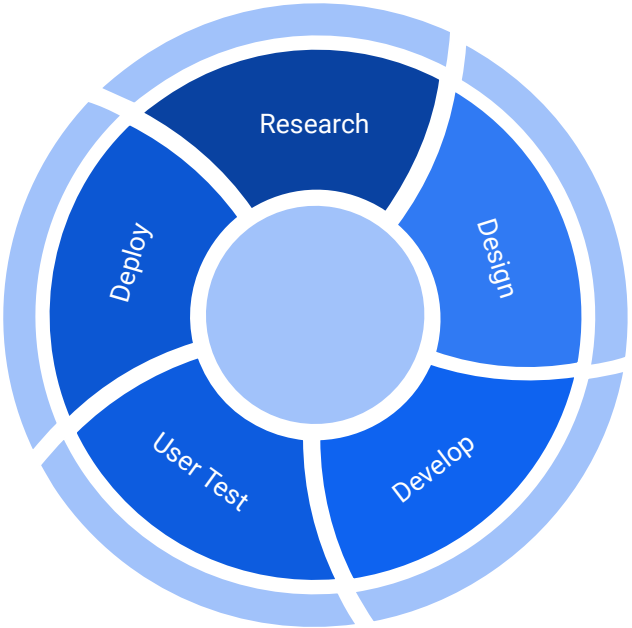
However, new schedules must abide by the mission's constraints and resources while not overwhelming astronauts.

Software tools must introduce more automation to everyday tasking:

- Automated constraint checking & verification.
- Suggested rescheduling.

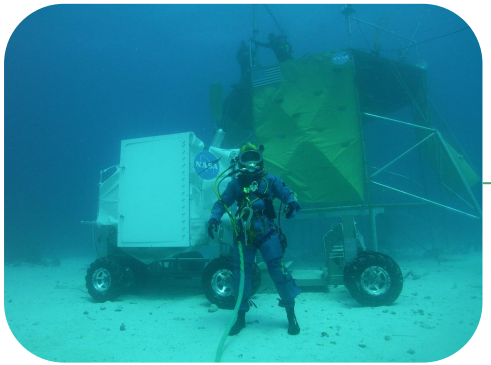
Human-Computer Interaction (HCI) Process

- Human-Computer Interaction (HCI) Group follows HCI principles.
- Since 2002, HCI group has deployed software tools that enable spaceflight operations.
- Emphasize human-centered design process that includes:
  - Observation of work in operations,
  - Usability testing,
  - Iterative deployment software processes.



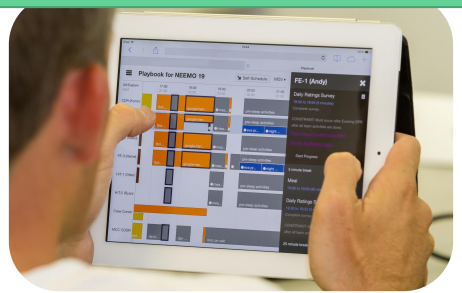
Combining Spaceflight & Analog Operational Experience

Mission Control supporting human spaceflight



Spaceflight Earth analog operations

deploying & testing software in analogs



Playbook: self-scheduling & execution software

Designed for tablet

Drag / drop for touch

Collaborative

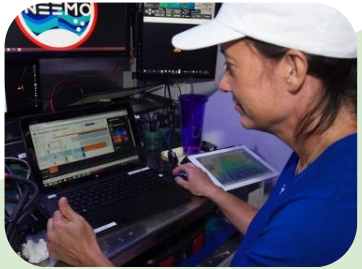
Schedule for multiple crew members

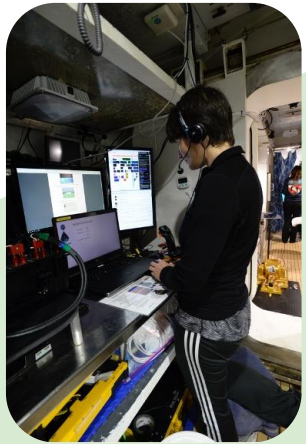
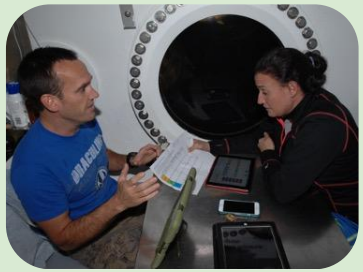
Incremental Evaluations

Playbook







self-scheduling tasks that were varied, differing complexity, and novel strategies

Subjective Feedback

“Three things you liked about Playbook?”

Over ten years, simple word frequency count indicated that ~15% of positive comments related to usability.

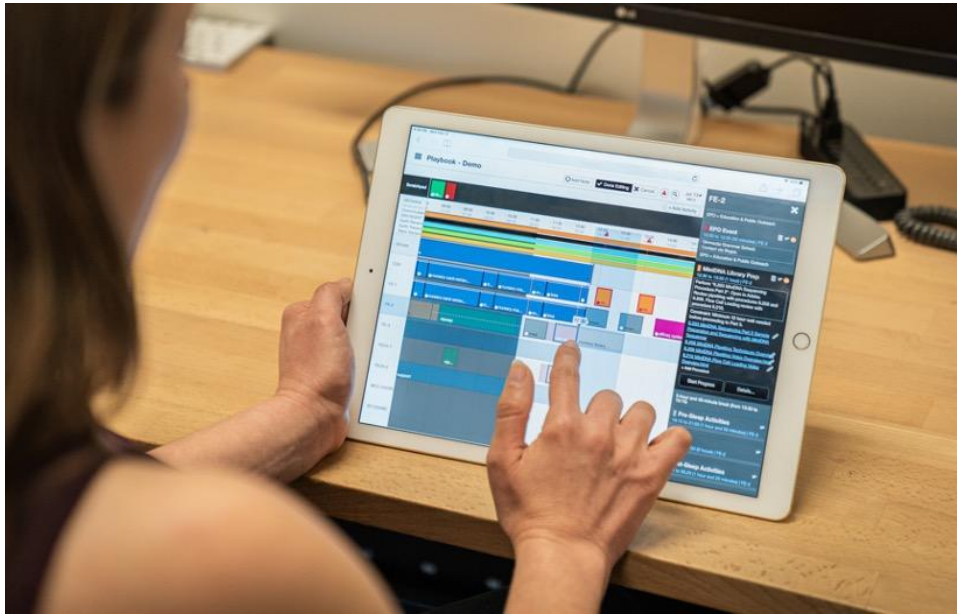
Subjective feedback identified that our tool has high usability.

“*Simple, convenient and intuitive design*” — NEEMO crew (2014)

“*Conveys lots of information concisely*” – NEEMO crew (2017)

“*Intuitive*” – NEEMO crew (2019)

Controlled Laboratory Study



Quantifying self-scheduling performance as a function of scheduling task complexity.

- Controlled lab experiment to measure task effectiveness, efficiency, workload, situation awareness, **trust**, and **usability**.

Experiment Design

Thirty-one participants in a human-in-the-loop experiment used our self-scheduling software tool, Playbook.

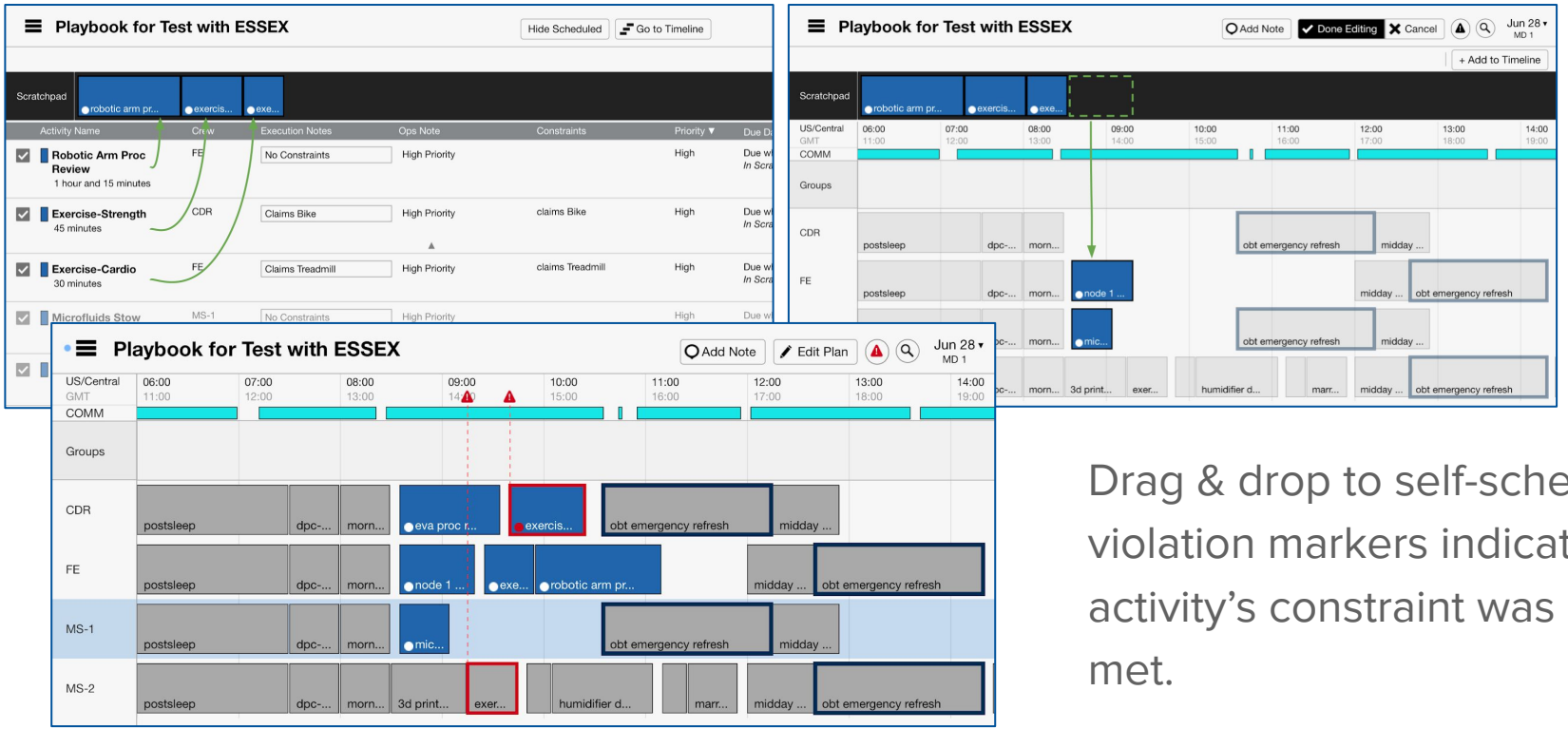
- Participants trained and completed nine trials, including a baseline trial.
- Participants completed self-scheduling tasks, creating a timeline that did not have any violations.

Dependent Variables

- *Efficiency & Effectiveness Performance*
- *Situation Awareness*
- *Workload*
- Trust
- Usability

Trust & usability were measured at the end of the experiment.

Self-Scheduling Task



Drag & drop to self-schedule; violation markers indicated if activity's constraint was not met.

Results

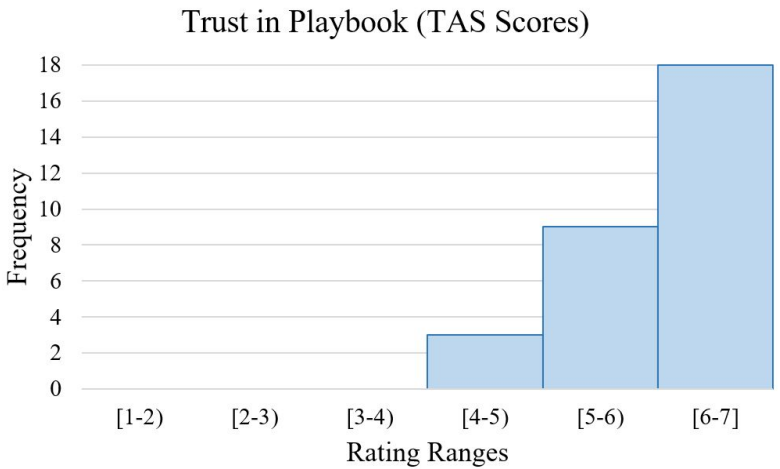
Trust

Trust was measured with the Trust in Automated Systems (TAS) Scale survey.

- Jian et al., 1998.

Participants reported relatively high trust in Playbook.

- Average trust score: 5.94 (SD = 0.64)



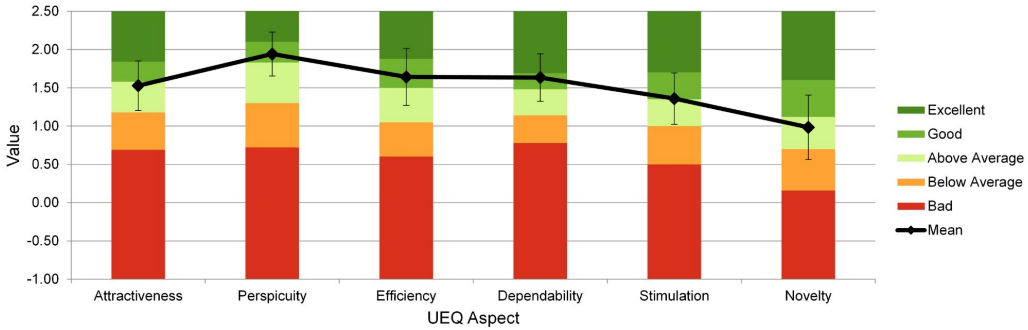
Results

Usability

User Experience Questionnaire (UEQ) characterizes usability across six factors (attractiveness, perspicuity, efficiency, dependability, stimulation, and novelty)

- Schrepp, Hinderks, & Thomaschewski, 2014

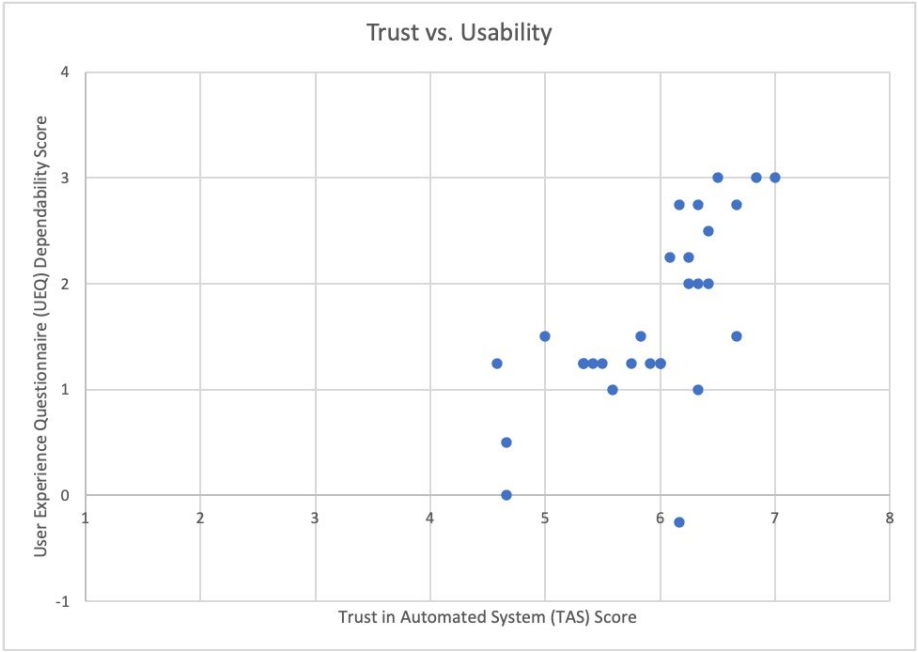
Playbook's usability was evaluated positively on most UEQ metrics.



Correlations Between Usability & Trust

Significant correlations between trust and all of the UEQ metrics.

- Strong correlation ($r > 0.5$) with UEQ
  - attractiveness, $p < 0.001$;
  - efficiency, $p < 0.001$;
  - dependability, $p < 0.001$;
  - stimulation, $p < 0.001$; and
  - novelty, $p < 0.01$.
- Moderate correlation between trust and perspicuity, $r(27) = 0.41$, $p < 0.01$.



Discussion

We confirmed subjective feedback about Playbook's high level of usability as well as found it correlated with learned trust.

Learned trust (Hoff & Bashir, 2015) develops from a user's (past or current) interactions with a system, which are driven by system performance and design features.

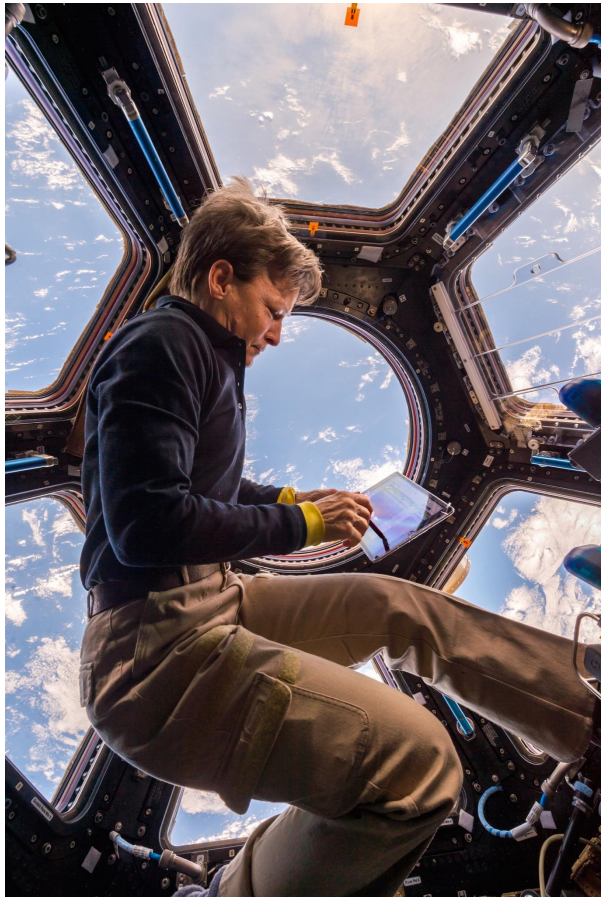
- Since participants did not have previous experience with Playbook, this trust is based on use during the experiment.

We did not find correlations between trust and other measures (workload, covariates such as dispositional trust & experience in scheduling).

Future work: collecting trust & usability within analog setting.

Concluding Remarks

- We identified a strong relationship between trust and usability.
- Other research has found a relationship between these factors. However, within the space domain, we have not identified any study that directly explored this relationship.
- Findings suggest that highly usable advanced software enables trust in automated systems for future autonomous crew.



Backup Slides

Collaborators: John Karasinski, Ph.D. (NASA); Shivang Shelat (SJSURF); Jimin Zheng (SJSURF).

Relevant publications:

Shelat, S., Karasinski, J. A., Flynn-Evans, E. E., & Marquez, J. J. (2022, June). Evaluation of user experience of self-scheduling software for astronauts: Defining a satisfaction baseline. In International Conference on Human-Computer Interaction (pp. 433-445). Cham: Springer International Publishing.

Marquez, J. J., Sullivan, D., & Karasinski, J. (2022). The role of trust and usability to enable spaceflight crew autonomy. In ASCEND 2022 (p. 4262).

Marquez, J. J., Edwards, T., Karasinski, J. A., Lee, C. N., Shyr, M. C., Miller, C. L., & Brandt, S. L. (2023). Human performance of novice schedulers for complex spaceflight operations timelines. Human Factors, 65(6), 1183-1198.

What are NASA's Standards for future systems?

- From NASA Spaceflight Human-System Standard 3001, Vol. 2 (Rev. D) - Human Factors, Habitability, and Environmental Health (2023):
  - Standards emphasize Human-Centered Task Analysis, Human Error Analysis, and Iterative Developmental Testing.
  - Standards also include requirements for Operability (effectiveness & efficiency), Usability, and Situation Awareness, while minimizing Errors and excessive Workload.