Crew Autonomy Scheduling: Scheduling Performance Pilot Study

Andrew Torr, Candice Lee, and Jessica Marquez, Ph.D.
NASA Ames Research Center (ARC)

Objective & Background

The purpose of this pilot study is to quantify crew performance in self-scheduling through Playbook, a mobile-based scheduling and planning tool. By investigating human performance within the task of self-scheduling, we can further develop countermeasures that can mitigate deficient scheduling performance, and evaluate changes as a result of these countermeasures. Moreover, this research can advise the development of standards and guidelines for autonomous crews in future missions.

In human spaceflight today, the task of planning crew member's schedules falls to Ops Planners. It takes many weeks to plan due to the complex impact each day-to-day activity can have on other activities, crew members, and resources. These impacts are measured as constraints, and these constraints can result in temporal, ordering, or resource requirements. As future spaceflight missions span longer distances and the latency of communication between the crew and Mission Control Center (MCC) increases, the need for crew members to work independently from MCC will also increase. This results in a need for crew members to be able to autonomously plan and adjust their own schedules.

Methods

- 11 Playbook-naive participants were recruited from NASA Ames to participate in the lab experiment and demographic information was collected.
- Participants were instructed how to use Playbook on an iPad and were given four planning problems for practice.
- A total of nine trials (one plan per trial) were presented along with a set of tasks to schedule. Immediately following each planning problem, the participant was prompted to answer one multiple choice question and to complete the NASA Task Load Index.
- All participants were presented with all conditions, in the same randomized order.
- After completing all nine trials, participants completed the System Trust Scale Survey, and the User Experience Questionnaire

References

[1] Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): results of empirical and theoretical research. *Advances in Psychology Human Mental Workload*, 139-183.

[2] Durso, F. T., Hackworth, C. A., Truitt, T. Crutchfield, J. and Manning, C. A. (1998) Situation awareness as a predictor of performance in en route air traffic controllers. *Air Traffic Quarterly 6*, 1-20.

[3] Jian, J., Bisantz, A., & Drury, C. (2000). Foundations for an empirically determined scale of trust in automated systems. *International Journal of Cognitive Ergonomics, 4:1*, 53-71.

[4] Laugwitz, Bettina & Held, Theo & Schrepp, Martin. (2008). Construction and evaluation of a user experience questionnaire. USAB 2008. 5298. 63-76.

Experimental Design

Independent Variables:

We established 3 levels for each of our variables (low, medium, high), resulting in a 3 x 3 Within-Subjects design, 9 conditions in total. One trial was created for each condition, and all participants completed all trials.

- IV 1: Number of activities (12, 24, 36)
- IV 2: Percentage of activities with unary temporal (start/end time) constraint (0%, 33%, 66%)

Dependent Variables:

We were interested in measuring human performance, behavior, and overall experience in scheduling tasks, including:

- Performance metrics:
  - Efficiency (trial duration and completion)
  - Effectiveness (unscheduled time in timeline)
  - Violations created and resolved
  - Amount & Priority of activities unscheduled
- Workload: NASA Task Load Index (NASA-TLX) [1]
- Situation Awareness: Situational Present Assessment Method (SPAM) [2]
- Trust: System Trust Scale (STS) [3]
- User Experience: User Experience Questionnaire (UEQ) [4]

Test Platform:

The experiment was conducted using Playbook, a web based crew planning and scheduling software tool designed to assist crew members in self-scheduling tasks.

Playbook has been tested extensively in a variety of NASA analog missions as the main operational planning tool for crew members to view and modify mission plans.

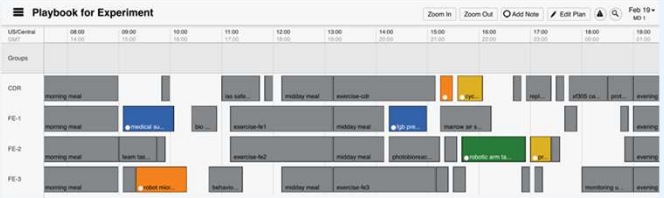

Figure 1. ***Playbook Timeline View*** - Screenshot of Playbook's timeline view. Grey tasks are inflexible and cannot be rescheduled. Colored tasks are flexible, and can be scheduled where time is available, as long as constraints do not create violations.

Results

Time on Task:

Mean time on task (the time for a subject to complete a plan) generally increased as the plan increased in total number of tasks and percentage of tasks with temporal constraints. The condition with 66% of 36 activities took the longest, with a mean time of 11 minutes and 59.9 seconds. The shortest mean time was 2 minutes and 19 seconds for the 12 activity condition with 0% constraints.

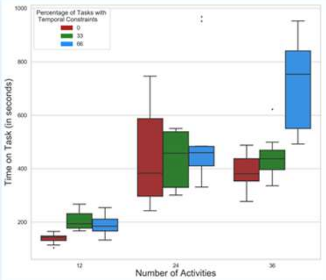

Figure 2. ***Time on Task*** - Graph shows the mean duration for each of the nine conditions.

User Experience Questionnaire (UEQ):

The UEQ revealed positive scores overall, with high ratings for the categories Perspicuity, Efficiency, and Dependability, which resulted in an overall Pragmatic Quality score of 1.55.

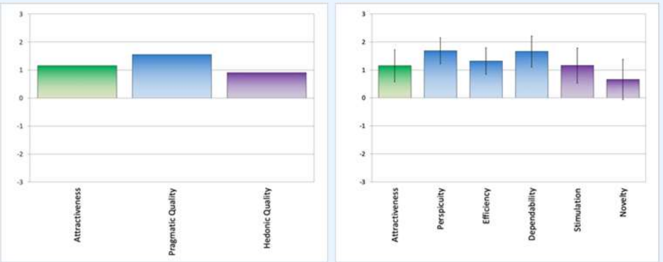

Figure 3. ***User Experience Questionnaire (UEQ)*** - The graph on the left shows combined scores for Attractiveness, Pragmatic Quality, and Hedonic Quality. The graph on the right shows scores for individual categories.

Future Work

- Analysis of performance data including measuring and comparing unscheduled time remaining, violations created and resolved, and number & priority of tasks left unscheduled.
- Analysis of NASA-TLX and SPAM data to measure the effect of the independent variables on Workload and Situation Awareness, and to help better understand the nature of plan complexity.
- Analysis of STS data, to better understand the level of trust for Playbook and self scheduling tools.