Management of Multiple Human Supervisory Control Tasks
Overview

- Introduction
- Motivation
- Background
- Simulation and Interface Design
- Hypotheses
- Methods
- Results and Discussion
- Conclusion
- Questions and Comments
Introduction – What is Human Supervisory Control (HSC)?

- **The five cyclical steps:**
  - **Planning** a computer-based task
  - **Communicating** to the computer what was planned
  - **Monitoring** the computer’s actions for errors and/or failures
  - **Intervening** when the plan has been completed or the computer requires assistance
  - The human and computer **learn** from the experience
Motivation

- **The role of multiple HSC tasks in network-centric operations**
  - HSC tasks are primarily cognitive in nature and generally do not require constant attention and/or control

- **Multiple HSC application of particular interest to military planners**
  - Operator supervision of a swarm of unmanned vehicles
    - Want to maximize the number of unmanned vehicles a single operator can supervise, under time pressure
    - Major limiting factor is operator workload, which could drive system wait times

- **The role of automated decision support**
  - Not clear what type or level of decision support is appropriate for supervision of multiple vehicles
**Background – Human Supervision of Multiple Vehicles**

- A framework for estimating the number of vehicles one human could control was proposed by Olsen and Wood (2004) with regard to traditional human-robot interactions.

- **Two main concepts in original work**
  - Interaction Time (IT)
  - Neglect Time (NT)

- **Fanout equation**
  - \# vehicles that a single operator can supervise = NT/IT + 1

- **Missing an essential element**
  - Wait Time (WT)
    - Dramatically impacts system performance and risk of failure in time-critical applications (e.g. C²)
Background – The Relationship Between IT, NT, and WT
• Three main categories of wait times

1. Interaction Wait Time (WTI)
   - Interaction times that occur while the vehicle is in a degraded state

2. Wait Time in the Queue (WTQ)
   - Result from vehicles requiring attention simultaneously or near simultaneously

3. Situation Awareness Wait Time (WTSA)
   - Result from loss of situational awareness, operator does not realize vehicle is waiting
The three major screen elements on the first display are:

1. **Map Display**
2. **Mission Time**
3. **Mission Planning & Execution**
The four major screen elements on the second display are:

1. UAV Status
2. Chat Box
3. UAV Health & Status Updates
4. Decision Support

**Simulation and Interface Design – Decision Support Display**

### MISSION STATUS

<table>
<thead>
<tr>
<th>Status</th>
<th>Current Target</th>
<th>Current</th>
<th>Course</th>
<th>Speed</th>
<th>Payload Ready</th>
</tr>
</thead>
<tbody>
<tr>
<td>GD3</td>
<td>T-14H</td>
<td>26°54’49” N</td>
<td>180°</td>
<td>2500 m</td>
<td></td>
</tr>
<tr>
<td>GD1</td>
<td>T-17L</td>
<td>26°54’49” N</td>
<td>2500 m</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GD2</td>
<td>T-10L</td>
<td>26°54’49” N</td>
<td>2500 m</td>
<td></td>
<td></td>
</tr>
<tr>
<td>GD3</td>
<td>T-11L</td>
<td>26°54’49” N</td>
<td>2500 m</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Message History**

- Intelligence (12:05:25): The following targets now require NDA T-41.
  - Base (12:05:25): Which UAVs’ selection plans would have to be revised if threat H-2 moved immediately south-west (SW) of Base?
  - Operator (12:05:25): 4
  - Base (12:06:36): Request to move TOT for T-41L decayed.
  - Base (12:07:00): Request to delay TOT for T-41L decayed.
  - Base (12:07:00): Request to delay TOT for T-41L decayed.

**Health & Status Updates**

- UAV 3 (12:04:25) → Available to fire for target T-16M
- UAV 3 (12:04:55) → Firing on T-16M
- UAV 3 (12:05:00) → Firing completed for target T-16M
- UAV 3 (12:05:45) → Available to arm for target T-4L
- UAV 3 (12:05:45) → Arming for T-4L
- UAV 4 (12:05:00) → Available to fire for target T-4L
- UAV 4 (12:05:00) → Firing on T-4L
- UAV 4 (12:06:00) → Firing completed for target T-4L
- UAV 1 (12:06:30) → Available to arm for target T-41H
- UAV 1 (12:06:30) → Arming for T-41H
- UAV 1 (12:06:30) → Available to fire for target T-41H
- UAV 1 (12:06:30) → Firing on T-41H
- UAV 1 (12:06:30) → Firing completed for target T-41H
Methods – Experimental Apparatus
**Methods – Experimental Design – Independent Variables**

**Primary: Level of Decision Support (Scheduling Assistance)**
- Between subjects
- Four levels
  1. Manual
  2. Passive
  3. Active
  4. Super Active

**Secondary: Amount of Schedule Re-planning (Operational Tempo)**
- Within subjects
- Two levels, high and low
**Methods** – Experimental Design – Types of Decision Support

### Manual

<table>
<thead>
<tr>
<th>MAV</th>
<th>Current Target</th>
<th>New Waypoint or Landingpoint</th>
<th>New Expected Action</th>
<th>Upcoming Action Targets</th>
<th>Measure From</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>T-6M</td>
<td></td>
<td></td>
<td>T-6H</td>
<td>10/25/155</td>
</tr>
<tr>
<td>2</td>
<td>T-10M</td>
<td></td>
<td></td>
<td>T-8L</td>
<td>10/25/155</td>
</tr>
<tr>
<td>3</td>
<td>T-12M</td>
<td></td>
<td></td>
<td>T-14L</td>
<td>10/25/155</td>
</tr>
<tr>
<td>4</td>
<td>T-14M</td>
<td></td>
<td></td>
<td>T-8L</td>
<td>10/25/155</td>
</tr>
</tbody>
</table>

### Passive

<table>
<thead>
<tr>
<th>MAV</th>
<th>Current Target</th>
<th>New Waypoint or Landingpoint</th>
<th>New Expected Action</th>
<th>Upcoming Action Targets</th>
<th>Measure From</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>T-6M</td>
<td></td>
<td></td>
<td>T-6H</td>
<td>10/25/155</td>
</tr>
<tr>
<td>2</td>
<td>T-10M</td>
<td></td>
<td></td>
<td>T-8L</td>
<td>10/25/155</td>
</tr>
<tr>
<td>3</td>
<td>T-12M</td>
<td></td>
<td></td>
<td>T-14L</td>
<td>10/25/155</td>
</tr>
<tr>
<td>4</td>
<td>T-14M</td>
<td></td>
<td></td>
<td>T-8L</td>
<td>10/25/155</td>
</tr>
</tbody>
</table>

### Active

<table>
<thead>
<tr>
<th>MAV</th>
<th>Current Target</th>
<th>New Waypoint or Landingpoint</th>
<th>New Expected Action</th>
<th>Upcoming Action Targets</th>
<th>Measure From</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>T-6M</td>
<td></td>
<td></td>
<td>T-6H</td>
<td>10/25/155</td>
</tr>
<tr>
<td>2</td>
<td>T-10M</td>
<td></td>
<td></td>
<td>T-8L</td>
<td>10/25/155</td>
</tr>
<tr>
<td>3</td>
<td>T-12M</td>
<td></td>
<td></td>
<td>T-14L</td>
<td>10/25/155</td>
</tr>
<tr>
<td>4</td>
<td>T-14M</td>
<td></td>
<td></td>
<td>T-8L</td>
<td>10/25/155</td>
</tr>
</tbody>
</table>

### Super Active

<table>
<thead>
<tr>
<th>MAV</th>
<th>Current Target</th>
<th>New Waypoint or Landingpoint</th>
<th>New Expected Action</th>
<th>Upcoming Action Targets</th>
<th>Measure From</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>T-6M</td>
<td></td>
<td></td>
<td>T-6H</td>
<td>10/25/155</td>
</tr>
<tr>
<td>2</td>
<td>T-10M</td>
<td></td>
<td></td>
<td>T-8L</td>
<td>10/25/155</td>
</tr>
<tr>
<td>3</td>
<td>T-12M</td>
<td></td>
<td></td>
<td>T-14L</td>
<td>10/25/155</td>
</tr>
<tr>
<td>4</td>
<td>T-14M</td>
<td></td>
<td></td>
<td>T-8L</td>
<td>10/25/155</td>
</tr>
</tbody>
</table>
Hypotheses – Wait Times

- **WTI**
  - Should decrease with higher levels of automation
    - Visual timeline in all but manual level allows users to more easily determine inter-vehicle relationships
    - Recommendations in active level decrease planning time
    - Super active level eliminates execution time

- **WTQ**
  - Follows same trend as WTI

- **WTSA**
  - Situation awareness can decrease under both high and low workloads
Methods – Experimental Design

- **Human Performance Measures**
  - Performance Score
  - Number of Time on Target Delays Requested

- **System Performance Measures**
  - Wait Times
    - WTI, WTQ, WTSA

- **Situation Awareness**
  - Quasi-objective SA score on a 0-5 scale
    - Objective rating scales, but subjectively chosen by expert observer

- **Critical Events**
  - Number of times targets were *incorrectly* destroyed
Results and Discussion – Performance Score

Level of re-planning significant \((p = 0.001)\)

Level of automation marginally significant \((p = 0.076)\)

- **Low re-planning**
  - No difference between automation levels

- **High re-planning**
  - Super active had significantly higher performance scores than active \((p = 0.032)\)
  - Poor performance of the active level is surprising in light of previous studies
Results and Discussion – Time on Target Delays

Level of re-planning marginally significant ($p = 0.059$)

Level of automation marginally significant ($p = 0.096$)

- High re-planning
  - TOT requests under active level significantly higher ($p = 0.065$)
  - The driver for poor performance in the active level
  - Subjects in the active level tried to globally optimize their schedules and generally failed
Results and Discussion – Wait Time Interaction (WTI)

Level of re-planning not significant (p = 0.170)

Level of automation significant (p = 0.003)

- Significantly less total WTI for super active and active levels than manual and passive levels
- Reasonably consistent with expectations
**Results and Discussion – Wait Time in the Queue (WTQ)**

**Level of re-planning significant**

(p = 0.001)

**Level of automation marginally significant**

(p = 0.063)

- Consistent with expectations, except:
  - Under high re-planning, active level significantly higher WTQ than super active or passive, same as manual
  - Task queues built significantly higher in the active condition when under high workload
  - Due to extra time operators spent trying to adjust schedule
Results and Discussion – Situation Awareness Wait Time (WTSA)

Level of re-planning significant
(p = 0.001)

Level of automation not significant
(p = 0.144)

- Super active significantly less WTSA than manual and active across both re-planning conditions
- Results not consistent with expectations
  - Inappropriate fixation on visual timeline
  - Further problems under active as users tried to globally optimize
Results and Discussion – Situation Awareness (SA)

- **Overall**
  - Level of automation not significant (p=0.112)
    - Same trends as WTSA
- **Breakdown**
  - SA is generally thought to have three levels (Endsley, 1995):
    1. Perception of important environmental cues
    2. Comprehension of the situation
    3. Projection of future events and dynamics
  - Super active had significantly higher level 2 SA
    - More time to observe events on displays
    - No improvement in level three SA
Results and Discussion – Wait Time Proportions

- Total system wait time dominated by WTSA regardless of level of workload
  - WTI, WTQ can be reduced by greater autonomy
  - WTSA, which dominates total wait time, cannot be completely eliminated

![Pie charts showing wait time proportions for low and high re-planning scenarios.](image)
Results and Discussion – Critical Events

- Occurred when operators erroneously fired upon targets
  - Friendly fire incidents, etc.
- Virtually no critical errors under low tempo condition
- Results suggest operators under manual and super active made more critical errors under high workload
Conclusion

- Super Active level of automation had the best human and system performance, but a higher number of catastrophic events
- Active level of automation unexpectedly produced the worst performance
  - Automation caused operators to attempt to globally optimize schedule, overwhelming them
- Passive level of automation best overall cost and performance benefits
  - A solid performer with no major drawbacks
- Total system wait time was dominated by wait time caused by lack of situation awareness
- Predictive model was good for WTI, WTQ, but not WTSA
  - Operators were never “under-loaded” due to number of vehicles being supervised
Questions or Comments?

MIT Humans and Automation Lab website

http://halab.mit.edu

Primary investigators’ contact information

pmitchel@mit.edu

missyc@mit.edu
References


**Background – Levels of Automation**

<table>
<thead>
<tr>
<th>Automation Level</th>
<th>Automation Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>The computer offers no assistance: human must take all decision and actions.</td>
</tr>
<tr>
<td>2</td>
<td>The computer offers a complete set of decision/action alternatives, or</td>
</tr>
<tr>
<td>3</td>
<td>narrows the selection down to a few, or</td>
</tr>
<tr>
<td>4</td>
<td>suggests one alternative, and</td>
</tr>
<tr>
<td>5</td>
<td>executes that suggestion if the human approves, or</td>
</tr>
<tr>
<td>6</td>
<td>allows the human a restricted time to veto before automatic execution, or</td>
</tr>
<tr>
<td>7</td>
<td>executes automatically, then necessarily informs humans, and</td>
</tr>
<tr>
<td>8</td>
<td>informs the human only if asked, or</td>
</tr>
<tr>
<td>9</td>
<td>informs the human only if it, the computer, decides to.</td>
</tr>
<tr>
<td>10</td>
<td>The computer decides everything and acts autonomously, ignoring the human.</td>
</tr>
</tbody>
</table>
12 subjects total, 3 per automation level
- Combination of MIT students, ROTC, and active duty military personnel
- Average age = 26.3 years, 10 male, 2 female
- 9 are ROTC/Air Force officers
  - Mostly 2nd Lieutenants but up to Lieutenant Colonel
  - 1-20 years experience, median 3
- 9 are pilots
  - Average flight hours 120
- 2 had previous (small) UAV experience
Results and Discussion – Total System Wait Time (WTT)

Level of re-planning significant (p < 0.001)

Level of automation significant (p = 0.018)

- Super active WTT significantly less than all other automation levels
- High re-planning only
  - Active level significantly higher than super active
Results and Discussion – Wait Time Interaction (WTI)

Level of re-planning not significant \( (p = 0.170) \)

Level of automation significant \( (p = 0.003) \)
- Significantly less for super active and active levels than manual and passive levels
Results and Discussion – Wait Time in the Queue (WTQ)

Level of re-planning significant (p = 0.001)

Level of automation marginally significant (p = 0.063)

- High re-planning only
  - Active level significantly higher WTQ than super active or passive, same as manual
**Results and Discussion – Situation Awareness Wait Time (WTSA)**

**Level of re-planning significant (p = 0.001)**

**Level of automation not significant (p = 0.144)**

- Cell comparisons
  - Super active significantly less than manual and active across both conditions
  - HRP, active significantly higher than manual and super active
Results and Discussion – Subjective Workload

Level of re-planning significant (p < 0.001)
• No significant difference in subjective workload for the manual level under different operational tempos

Level of automation not significant (p = 0.779)
Results and Discussion – Secondary Workload (Spare Capacity)

Level of re-planning significant (p=0.003)
- Spare capacity of the manual and passive levels did not change across re-planning conditions

Level of automation significant (p=0.002)
- Manual and super-active automation levels statistically have the same spare capacity, which is lower than active and passive levels
  - Manual = workload
  - Super Active = low SA
## Results and Discussion – Best and Worst Performers

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>Level of Automation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Worst Performer</td>
</tr>
<tr>
<td>Performance Score</td>
<td>Active</td>
</tr>
<tr>
<td>Number of TOT Delay Requests</td>
<td>Active</td>
</tr>
<tr>
<td>Interaction Wait Time (WTI)</td>
<td>Manual / Passive</td>
</tr>
<tr>
<td>Wait Time in the Queue (WTQ)</td>
<td>Active / Manual</td>
</tr>
<tr>
<td>Situation Awareness Wait Time (WTSA)</td>
<td>Active</td>
</tr>
<tr>
<td>Situation Awareness (SA)</td>
<td>Active</td>
</tr>
<tr>
<td>Erroneous Target Critical Events</td>
<td>Super Active / Manual</td>
</tr>
<tr>
<td>Subjective Workload</td>
<td>Active / Manual</td>
</tr>
<tr>
<td>Secondary Workload</td>
<td>Manual</td>
</tr>
</tbody>
</table>
## Results and Discussion – Summary of Main Effects (p-values)

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>Independent Variable</th>
<th>Level of Automation</th>
<th>Level of Re-planning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Performance Score</td>
<td></td>
<td>0.076</td>
<td>0.001</td>
</tr>
<tr>
<td>Number of TOT Delay Requests</td>
<td></td>
<td>0.096</td>
<td>0.059</td>
</tr>
<tr>
<td>Interaction Wait Time (WTI)</td>
<td></td>
<td>0.003</td>
<td>0.170</td>
</tr>
<tr>
<td>Wait Time in the Queue (WTQ)</td>
<td></td>
<td>0.063</td>
<td>0.001</td>
</tr>
<tr>
<td>Situation Awareness Wait Time (WTSA)</td>
<td></td>
<td>0.144</td>
<td>0.001</td>
</tr>
<tr>
<td>Situation Awareness (SA)</td>
<td></td>
<td>0.112</td>
<td>0.002</td>
</tr>
<tr>
<td>Erroneous Target Critical Events</td>
<td></td>
<td>0.878</td>
<td>0.034</td>
</tr>
<tr>
<td>Subjective Workload</td>
<td></td>
<td>0.779</td>
<td>&lt;0.001</td>
</tr>
<tr>
<td>Secondary Workload</td>
<td></td>
<td>0.002</td>
<td>0.003</td>
</tr>
</tbody>
</table>

**LEGEND**
(a=0.05)
- Significant Result
- Marginally Significant Result
Results and Discussion – Overloaded Operator Coping Strategies

Cognitive Shedding

Degraded Level of Management
Results and Discussion – Example Test Session – 1

Super-Active Automation
High Re-planning Scenario

12:00:59
Super-Active Automation
High Re-planning Scenario

12:02:19
Results and Discussion – Example Test Session – 3

Super-Active Automation
High Re-planning Scenario

12:05:34
Results and Discussion – Example Test Session – 4

Super-Active Automation
High Re-planning Scenario

12:07:39
Results and Discussion – Example Test Session – 5

Super-Active Automation
High Re-planning Scenario

12:08:54
Results and Discussion – Example Test Session – 6

Super-Active Automation
High Re-planning Scenario

12:10:29
Results and Discussion – Example Test Session – 7

Super-Active Automation
High Re-planning Scenario

12:12:44
Results and Discussion

Super-Active Automation
High Re-planning Scenario

12:15:24
Results and Discussion – Example Test Session – 9

Super-Active Automation
High Re-planning Scenario

12:17:14
Super-Active Automation
High Re-planning Scenario

12:20:09
Results and Discussion – Example Test Session – 11

Super-Active Automation
High Re-planning Scenario

12:21:49
Results and Discussion – Example Test Session – 12

Super-Active Automation
High Re-planning Scenario

12:23:04