Introduction to Human Factors

- Protocols of human testing
- "Human Factors" - what is it?
- Methodologies for understanding human activities
- Samples from past and ongoing SSL research

© 2019 David L. Akin - All rights reserved
http://spacecraft.ssl.umd.edu
Testing Humans

• Not like testing hardware
  – Individual variability
  – Alertness and fatigue
  – Motivational differences
  – Skill differences

• Focus on obtaining statistically significant sample size

• Investigator responsible for ethical treatment of test subjects
Ethics of Human Testing - Safety

- Institutional Review Board (IRB) approval required for experiments using human subjects
- Legal and ethical responsibility to ensure informed consent is obtained and documented
- Subjects must be volunteers without being coerced (overt or covert) and free to withdraw at any time
- Subjects must be drawn from the most diverse pool possible within limitations of test requirements
Ethics of Human Testing - Diversity

• Federally supported research mandates the use of the most diverse test population possible
  – Gender
  – Race
  – Age

• Most space human testing is severely limited in feasible subject pool, number of available test subjects

• Ideal test pool ~40-60 test subjects; aerospace testing frequently limited to ~8-10 or even less
Ethics of Human Testing - Privacy

- Subjects must never be identified by name, initials, or other means which would allow identification of individuals.
- Subjects typically identified by numbers, letters, or other random labels.
- Lead investigators maintain index from code to test subject names under lock and key.
- Characteristics potentially relevant to performance (size, gender, experience, etc.) must be correlated to identifiers.
Collection of Subject Metrics

- Some parameters are potentially relevant to performance of human tasks
  - Gender
  - Physical size
  - Age
  - Relevant experience
  - etc.

- Data collected via questionnaire filled out prior to testing, investigator measurements, qualifying tests, etc.
Control of Variables - Motivation

• Performance is generally strongly driven by individual motivation (perfection, competition, boredom, etc.)

• Test subjects should not be exposed to other test subjects to avoid competition (particularly powerful in college-age males)

• Investigators generally create a set of instructions provided to test subject (often read aloud) with very specific instructions on desired parameters of testing (e.g., “minimum time”, “maximum precision”)
Control of Variables - Performance

• Test subject access to performance metrics is very highly restricted
  – Never provide comparative performance with other test subjects
  – Frequently do not tell subjects their own performance metrics after individual runs

• Subjects should seldom be allowed to observe other subjects performing task
  – Avoid adopting techniques from other subjects
  – Have to be concerned about subject-subject “hints”
Control of Variables - Learning

- People tend to improve with repetition
- Tends to follow exponential heuristic:
  \[ t_n = t_1 n^p \]
- Typically quantify learning as reduction following doubling of number of trials - "80%" where \( t_2 = 0.8t_1 \), \( t_4 = 0.8t_2 \), etc.

\[ LC\% = \frac{t_2}{t_1} = 2^p \implies p = \frac{\ln (LC\%)}{\ln (2)} \]

- Have to ensure test subjects have essentially leveled off on learning curve for task performance
Control of Variables - Trial Order

- People tend to prefer system they first train on
- Greatest number of repetitions will be on first trial configuration
- Where possible, have sufficient test population to have multiple subjects in each of all possible permutations of order for test matrix
- Where fully populated test matrix is not possible, at least have distribution in starting mode
The Scope of Human Factors

from Chapanis, Human Factors in Systems Engineering - Wiley Interscience, 1996
Subsumed within “Human Factors”

- Anthropometrics
- Bioastronautics
- Biomechanics
- Environmental medicine
- Ergonomics
- Experimental psychology
- Human engineering
- Human-machine systems design
- Human-computer interaction (HCI)
- Life sciences engineering
Introduction to Human Testing

UNIVERSITY OF MARYLAND
ENAE 697 - Space Human Factors and Life Support

Taxonomy of Human Factors Methods

- Data collection techniques
- Task analysis techniques
- Cognitive task analysis techniques
- Charting techniques
- Human error identification (HEI) techniques
- Mental workload assessment techniques
- Situational awareness measurement techniques
- Interface analysis techniques
- Systems design techniques
- Performance time prediction/assessment tech.
- Team performance analysis techniques
Data Collection Techniques

- Interviews
- Questionnaires
- Observation
Interviews

- Structured, semi-structured, unstructured
- Closed, open, and probing questions
- Focus groups instead of multiple interviews
- Interviews take the longest to train of all data collection techniques
- Should last a minimum of 20 and maximum of 40 minutes
- Data collection: notes, audio/video recorder
Questionnaires

- Types of questions used
  - Multiple choice
  - Rating scales ("strongly agree", "agree"...)
  - Paired associates ("Which is more difficult, A or B?")
  - Ranking ("On a scale of 1-10...")
  - Open-ended ("What did you think of...")
  - Closed questions ("yes/no")

- Easy to collate and reduce data
- Usually have poor voluntary response
- Needs to be designed well for best results
Analytical Hierarchy Process

- Considering a range of options, e.g., ice cream
  - Vanilla (V)
  - Peach (P)
  - Strawberry (S)
  - Chocolate (C)
- Could ask for a rank ordering, e.g. (1) vanilla, (2) strawberry, (3) peach, (4) chocolate - but that doesn’t give any information on how firm the rankings are
- Use pairwise comparisons to get numerical evaluation of the degree of preference
Pairwise Comparisons

• Ideally, do exhaustive combinations
  – Vanilla >> chocolate (strongly agree)
  – Vanilla >> peach (agree)
  – Vanilla >> strawberry (agree)
  – Peach >> chocolate (strongly agree)
  – Peach >> strawberry (disagree)
  – Strawberry >> chocolate (strongly agree)

• Number of required pairings out of N options is \((N)(N-1)/2\) - e.g., N=20 requires 190 pairings!

• Can use hierarchies of subgroupings to keep it manageable
Evaluation Metric

• Create a numerical scaling function, e.g.
  – “strongly agree” = 9
  – “agree” = 3
  – “neither agree nor disagree” = 1
  – “disagree” = 1/3
  – “strongly disagree” = 1/9

• Numerical rankings are arbitrary, but often follow geometric progressions
  – 9, 3, 1, 1/3, 1/9
  – 8, 4, 2, 1, 1/2, 1/4, 1/8
### Evaluation Matrix

- Fill out matrix preferring rows over columns

<table>
<thead>
<tr>
<th></th>
<th>C</th>
<th>S</th>
<th>P</th>
<th>V</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P</td>
<td>9</td>
<td>1/3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>V</td>
<td>9</td>
<td>3</td>
<td>3</td>
<td></td>
</tr>
</tbody>
</table>

- Fill out matrix preferring rows over columns
**Evaluation Matrix**

- Fill out matrix preferring rows over columns
- Fill opposite diagonal with reciprocals

<table>
<thead>
<tr>
<th></th>
<th>C</th>
<th>S</th>
<th>P</th>
<th>V</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>S</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P</td>
<td>9</td>
<td>1/3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>V</td>
<td>9</td>
<td>3</td>
<td>3</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>C</th>
<th>S</th>
<th>P</th>
<th>V</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td></td>
<td>1/9</td>
<td>1/9</td>
<td>1/9</td>
</tr>
<tr>
<td>S</td>
<td>9</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>P</td>
<td>9</td>
<td>1/3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>V</td>
<td>9</td>
<td>3</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**ENAE 697 - Space Human Factors and Life Support**
### Normalization of Matrix Elements

- Normalize columns by column sums

<table>
<thead>
<tr>
<th></th>
<th>C</th>
<th>S</th>
<th>P</th>
<th>V</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>1/9</td>
<td>1/9</td>
<td>1/9</td>
<td></td>
</tr>
<tr>
<td>S</td>
<td>9</td>
<td>3</td>
<td>1/3</td>
<td></td>
</tr>
<tr>
<td>P</td>
<td>9</td>
<td>1/3</td>
<td>1/3</td>
<td></td>
</tr>
<tr>
<td>V</td>
<td>9</td>
<td>3</td>
<td>3</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>C</th>
<th>S</th>
<th>P</th>
<th>V</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>0.032</td>
<td>0.018</td>
<td>0.143</td>
<td></td>
</tr>
<tr>
<td>S</td>
<td>0.333</td>
<td>0.491</td>
<td>0.429</td>
<td></td>
</tr>
<tr>
<td>P</td>
<td>0.333</td>
<td>0.097</td>
<td>0.429</td>
<td></td>
</tr>
<tr>
<td>V</td>
<td>0.333</td>
<td>0.871</td>
<td>0.491</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>C</th>
<th>S</th>
<th>P</th>
<th>V</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>27</td>
<td>3.44</td>
<td>6.11</td>
<td>0.78</td>
</tr>
</tbody>
</table>
### Evaluation of Hierarchy Among Options

- Average across the populated row elements

<table>
<thead>
<tr>
<th></th>
<th>C</th>
<th>S</th>
<th>P</th>
<th>V</th>
</tr>
</thead>
<tbody>
<tr>
<td>C</td>
<td>0.032</td>
<td>0.018</td>
<td>0.143</td>
<td>0.048</td>
</tr>
<tr>
<td>S</td>
<td>0.333</td>
<td>0.491</td>
<td>0.429</td>
<td>0.313</td>
</tr>
<tr>
<td>P</td>
<td>0.333</td>
<td>0.097</td>
<td>0.429</td>
<td>0.215</td>
</tr>
<tr>
<td>V</td>
<td>0.333</td>
<td>0.871</td>
<td>0.491</td>
<td>0.424 ← Top ranking</td>
</tr>
</tbody>
</table>
## Space Allocation and Crew Flow

<table>
<thead>
<tr>
<th>Unit</th>
<th>Description</th>
<th>% of habitable volume</th>
</tr>
</thead>
<tbody>
<tr>
<td>Work</td>
<td>Operational or Mission-related tasks</td>
<td>40%</td>
</tr>
<tr>
<td>Public</td>
<td>Dining, food, management, recreation, and exercise</td>
<td>25%</td>
</tr>
<tr>
<td>Personal</td>
<td>Sleeping, privacy, personal stowage</td>
<td>20%</td>
</tr>
<tr>
<td>Service</td>
<td>Hygiene, waste management, public stowage</td>
<td>15%</td>
</tr>
</tbody>
</table>

Data from Parker & Every (1972) and Schowalter & Malone (1972)
Analytical Hierarchy Process

- Used an Analytical Hierarchy Process survey to determine the relative importance of possible habitat functions for an MFH
  - Life support assumed present
  - Two-level AHP ranks 34 functions based on 90 pair-wise rankings
- Targeted population with experience in remote/confined environments:
  - Astronauts
  - Submariners/ship crews
    - “Submarines were found to be most similar overall to the space ship situation…”
      Source: Habitability Issues in Long-Duration Undersea and Space Missions Jul 1972
  - Arctic/Antarctic research scientists
    - “The south pole is the closest place to space on earth where a permanent, manned US presence exists, and represents a good scientific/logistics/operations analogue for future moon/mars missions”
      Source: Antarctic Exploration: Proxy for Safe, Sustainable Exploration of the Moon and Mars
Survey Hierarchy

- **Top-level matrix**
  - **Health and Hygiene**
    - Habitat cleanliness
    - Personal hygiene
    - Comfort of bathroom
    - Quality of medical
    - Clothing cleanliness
  - **Communications**
    - Comms functions
    - Comms time/day
    - Comms privacy
    - Connection quality
  - **General environ.**
    - Lighting control
    - Noise control
    - Windows
    - Odor control
    - Temp. control
    - Standing clearance
    - Food quality
  - **Recreation area**
    - Rec. time/day
    - Rec. space
    - Rec. variety
    - Rec. alone ratio
    - EVA time/day
    - Rec. privacy
  - **Exercise area**
    - Exe. time/day
    - Exe. space
    - Exe. variety
    - Exe. alone ratio
    - EVA time/day
    - Exe. privacy
  - **Work area**
    - Work time/day
    - Work space
    - Prep time/EVA
    - Work alone ratio
    - Work privacy
  - **Sleep area**
    - Sleep time/day
    - Sleep space
    - Sleep privacy
    - Sleep comfort
    - No hot-racking
Online AHP Survey
Data Analysis Method

- Subjective survey responses converted to numerical relative importance values and fed into AHP matrices:
  - “Much less important” = 0.125 ( = 2^{-3} )
  - “Moderately less important” = 0.354 ( = 2^{-1.5} )
  - “A little less important” = 0.707 ( = 2^{-0.5} )
  - “About as important” = 1.000 ( = 2^{0} )
  - “A little more important” = 1.414 ( = 2^{0.5} )
  - “Moderately more important” = 2.828 ( = 2^{1.5} )
  - “Much more important” = 8.000 ( = 2^{3} )

- Remaining matrix elements filled in with reciprocals of conjugate elements

- For each AHP matrix:
  - Importance values of each function or sub-category are the elements of the normalized principal eigenvector
  - “Consistency” is matrix size divided by the principal eigenvalue, with a value of 1 indicating complete consistency
  - Function importance values multiplied by importance value of the sub-category

- Overall importance values are the averaged values generated from all respondents, weighted by matrix consistency
AHP Results: Function Importance Values

Habitat functions

1. No hot racking
2. Quality of medical
3. Quality of comms
4. Personal hygiene
5. Work time/day
6. Work space
7. Sleep comfort
8. Prep time/EVA
9. Comms functions
10. Comms privacy
11. Cleanliness of habitat
12. Comfort of bathroom
13. Comms time/day
14. Sleep privacy
15. Food quality
16. Exercise variety
17. Cleanliness of clothing
18. Work alone ratio
19. Sleep time/day
20. Exercise space
21. Temperature control
22. Exercise time/day
23. Sleep space
24. Standing clearance
25. EVA time/day
26. Odor control
27. Exercise alone ratio
28. Noise control
29. Recreation space
30. Recreation variety
31. Recreation alone ratio
32. Recreation time/day
33. Lights control
34. Windows

Importance value
AHP Results: Important Functions

• Hot racking considered unacceptable, the most important function at 2.3 times the average importance value
• Medical facilities, communications connection quality, and personal hygiene round out vital functions
• Work time and space were highly ranked
• Non-physical recreation features considered especially unimportant
• Lighting quality and windows were the least important functions considered, with windows 0.35 times as important as the average function
• The most important function was 6.5 times as important as the least important function
AHP Results: Consistency and Variation

- Overall matrix consistency: 92.5%
  - Most consistent matrix: “Work space”, at 96.6%
  - Least consistent matrix: “General environmental quality”, at 90.3%
  - Importance value averages are weighted by matrix consistency to improve reliability of results

- Standard deviation and coefficient of variation were computed for each habitat function
  - Average standard deviation was 0.0215, average coefficient of variation was 73.4%
  - Greatest std. dev.: “No hot racking” (σ = .0645, \(c_v = 97.4\%\))
  - Greatest coeff. of variation: “Quality of comms” (σ = .0637, \(c_v = 112.5\%\))
  - Lowest std. dev.: “Recreation time per day” (σ = .0066, \(c_v = 45.5\%\))
AHP Results: Consistency

- Matrix size divided by principal eigenvalue is a measure of internal consistency, ranging from zero to one
  - Overall matrix consistency: 92.5%
  - Most consistent matrix: “Work space”, at 96.6%
  - Least consistent matrix: “General environmental quality”, at 90.3%

- Importance value averages are weighted by matrix consistency to improve reliability of results
  - Mean difference from un-weighted average: 1.7%
  - Greatest difference: recreation alone-time ratio (3.9% more important in weighted average)
**AHP: Demographics and Analysis of Variance**

Statistically significant variances, at 95% confidence

- Respondents:
  - By nationality:
    - American (15)
    - Italian (11)
    - French (2)
    - Romanian (1)
  - By experience:
    - Submarine (19)
    - Ship (11)
    - Arctic/Antarctic base (3)
    - Other (2)
  - By age group:
    - ≤40 years (16)
    - >40 years (13)

<table>
<thead>
<tr>
<th>Demographic set</th>
<th>Feature</th>
<th>Difference from complimentary set</th>
</tr>
</thead>
<tbody>
<tr>
<td>French</td>
<td>EVA time/day</td>
<td>+26.3%</td>
</tr>
<tr>
<td></td>
<td>Exercise alone ratio</td>
<td>-11.9%</td>
</tr>
<tr>
<td>American</td>
<td>Quality of comms</td>
<td>-5.3%</td>
</tr>
<tr>
<td>Ship crew members</td>
<td>Personal hygiene</td>
<td>+0.9%</td>
</tr>
<tr>
<td></td>
<td>Quality of medical</td>
<td>+1.3%</td>
</tr>
<tr>
<td></td>
<td>Recreation alone-time ratio</td>
<td>-0.8%</td>
</tr>
<tr>
<td></td>
<td>Sleep privacy</td>
<td>+0.3%</td>
</tr>
<tr>
<td>Submariners</td>
<td>Bathroom comfort</td>
<td>-0.6%</td>
</tr>
<tr>
<td>Age 40+</td>
<td>Comms privacy</td>
<td>-1.4%</td>
</tr>
<tr>
<td></td>
<td>Temperature control</td>
<td>-0.8%</td>
</tr>
</tbody>
</table>

- Performing ANOVA between astronaut and analogue populations can justify the statistical relevance of analogue populations
Fidelity of Analogue Environments

• The analogue environments considered in the survey may be of low fidelity, due to several factors:
  – Windows may be less important in environments with a static view/no external view
  – Affects of reduced gravity on the importance of habitat functions not accounted for
  – Ability to leave environment may impact importance of habitat functions

• Larger samples and samples of the astronaut population would be needed to identify statistical significance of variations between analogue and space environments
Observation

- Provides “real-life” insight into actual operations
- Intrusive (Heisenberg’s principle)
  - “fly on wall”
  - “hanging over the shoulder”
- Time consuming (1 hr of audio = 8 hrs of transcription)
- Difficult to set up, expensive, time consuming
- Generally requires teams to cover all critical areas
Task Analysis Techniques

- Hierarchical Task Analysis (HTA)
- Critical Path Analysis (CPA)
- Goals, Operators, and Selection Methods (GOMS)
- Verbal Protocol Analysis (VPA)
- Task Decomposition
- Sub Goal Template (SGT) approach
- Tabular Task Analysis
Hierarchical Task Analysis (HTA)

- Ubiquitous; perhaps most common TA method
- Easy to implement; minimal training
- Generic; can be applied to any domain at any desired level of depth
- Descriptive rather than analytical
- Laborious for complex domain tasks
- Simple flowcharting/outlining processes
HTA Graphical Example

from Stanton et.al., Human Factors Methods - Ashgate, 2005
<table>
<thead>
<tr>
<th>Step</th>
<th>Task Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.</td>
<td>Boil kettle</td>
</tr>
<tr>
<td>Plan 0:</td>
<td>Do 1 then 2 then 3 then 4 then 5</td>
</tr>
<tr>
<td>1.</td>
<td>Fill kettle</td>
</tr>
<tr>
<td>Plan 1:</td>
<td>Do 1 then 2 then 3 (if full then 4 else 3) then 5</td>
</tr>
<tr>
<td></td>
<td>Take to tap</td>
</tr>
<tr>
<td></td>
<td>Turn on water</td>
</tr>
<tr>
<td></td>
<td>Check level</td>
</tr>
<tr>
<td></td>
<td>Turn off water</td>
</tr>
<tr>
<td></td>
<td>Take to socket</td>
</tr>
<tr>
<td>2.</td>
<td>Switch kettle on</td>
</tr>
<tr>
<td>Plan 2:</td>
<td>Do 1 then 2</td>
</tr>
<tr>
<td></td>
<td>2.1 Plug into socket</td>
</tr>
<tr>
<td></td>
<td>2.2 Turn on power</td>
</tr>
<tr>
<td>3.</td>
<td>Check water in kettle</td>
</tr>
<tr>
<td>4.</td>
<td>Switch kettle off</td>
</tr>
<tr>
<td>5.</td>
<td>Pour water</td>
</tr>
<tr>
<td>Plan 5:</td>
<td>Do 1 then 2 then 3</td>
</tr>
<tr>
<td></td>
<td>5.1 Lift kettle</td>
</tr>
<tr>
<td></td>
<td>5.2 Direct spout</td>
</tr>
<tr>
<td></td>
<td>5.3 Til' kettle</td>
</tr>
<tr>
<td></td>
<td>5.4 Replace kettle</td>
</tr>
</tbody>
</table>

From Stanton et. al., Human Factors Methods - Ashgate, 2005
Cognitive Task Analysis Techniques

• Applied Cognitive Task Analysis (ACTA)
• Cognitive Walkthrough
• Critical Decision Method (CDM)
• Critical Incident Technique
Cognitive Walkthrough

- Select tasks to be analyzed
- Create task descriptions
- Determine the correct sequence of actions
- Identify the user population
- Describe the user’s initial goals
- Analyze the interaction between the user and the interfaces
  - Problems in selecting and evaluating an action
  - Changing goals due to execution and system response
Charting Techniques

- Process Charts
- Operational Sequence Diagrams
- Decision Action Diagram (DAD)
- Event Tree Analysis
- Fault Tree Analysis
- Murphy Diagrams
Event Tree Diagram

from Stanton et.al., Human Factors Methods - Ashgate, 2005
Fault Tree Analysis

from Stanton et.al., Human Factors Methods - Ashgate, 2005
Human Error Identification Techniques

- Cognitive Reliability Error Analysis Method (CREAM)
- Human Error Assessment and Reduction Technique (HEART)
- Human Error Identification in Systems Tool (HEIST)
- Human Error Template (HET)
- Human Error HAZOP
- Systematic Human Error Reduction and Prediction Approach (SHERPA)
HEI Techniques (continued)

• System for Predictive Error Analysis and Reduction (SPEAR)
• Task Analysis For Error Identification (TAFEI)
• Technique for Human Error Assessment (THEA)
• The HERA Framework
• Technique for the Retrospective and Predictive Analysis of Cognitive Errors in Air Traffic Control (TRACEr)
Human Error Template (HET)

- Standard taxonomy of errors
  - Failed to execute
  - Task execution incomplete
  - Task executed in wrong direction
  - Wrong task executed
  - Task repeated
  - Task executed on wrong component
  - Task executed too early/late/much/little
  - Misread information

- Apply and evaluate to each bottom-level task from HTA

- Consider consequence, criticality, interface role
Situational Awareness Measurements

- SA Requirements Analysis
- Situation Awareness Global Assessment Tech.
- Situation Awareness Rating Technique
- SA Subjective Workload Dominance Metric
- Situation Awareness Control Room Inventory
- Situation Awareness Rating Scales (SARS)
- Situation-Present Assessment Method (SPAM)
- SA Behavioral Rating Scales
- Mission Awareness Rating Scale (MARS)
- Crew Awareness Rating Scale (CARS)
Situational Awareness Rating Scales

• List of 31 criteria in 8 categories from ACM SA
  – General traits
  – Tactical game plan
  – System operation
  – Communication
  – Information interpretation
  – Tactical employment BVR
  – Tactical employment visual
  – Tactical employment general

• Each criteria self-rated on 1(poor)-6(best) scale
Mental Workload Assessment

- Primary Task Performance Measures
- Secondary Task Performance Measures
- Bedford Scale
- Defense Research Agency Workload Scale
- Instantaneous Self Assessment Workload
- Malvern Capacity Estimate
- Modified Cooper-Harper Rating
- NASA Task Load Index
- Subjective Workload Assessment Technique
- Workload Profile Technique
- Cognitive Task Load Analysis
Cooper-Harper Rating

Diagram:

- ADEQUACY FOR SELECTED TASK OR REQUIRED OPERATION
- AIRCRAFT CHARACTERISTICS
- DEMANDS ON THE PILOT IN SELECTED TASK OR REQUIRED OPERATION
- PILOT RATING

1. Excellent: Pilot compensation not a factor for desired performance
2. Highly desirable: Pilot compensation not a factor for desired performance
3. Good: Pilot compensation not a factor for desired performance
4. Negligible deficiencies: Minimal pilot compensation required for desired performance
5. Fair: Some mildly unpleasant deficiencies
6. Minor but annoying deficiencies
7. Moderately objectionable deficiencies
8. Very objectionable but tolerable deficiencies
9. Major Deficiencies
10. Definition of required operation involves designation of flight phase and/or subphases with accompanying conditions.

- Is it: satisfactory without improvement?
  - Yes
  - Deficiencies warrant improvement
    - Yes
    - Improvement mandatory
    - Major Deficiencies
      - Control will be lost during some portion of the required operation
  - No
  - Deficiencies require improvement
    - Minor but annoying deficiencies
    - Desired performance requires moderate pilot compensation
    - Moderately objectionable deficiencies
    - Adequate performance requires considerable pilot compensation
    - Very objectionable but tolerable deficiencies
    - Adequate performance requires extensive pilot compensation
    - Major Deficiencies
      - Intense pilot compensation is required to retain control
    - Major Deficiencies
      - Control will be lost during some portion of the required operation
Cooper-Harper Rating (close-up 1)
Cooper-Harper Rating (close-up 2)
Cooper-Harper Definition of Terms

DEFINITIONS FROM TN-D-5153

COMPENSATION
The measure of additional pilot effort and attention required to maintain a given level of performance in the face of deficient vehicle characteristics.

PERFORMANCE
The precision of control with respect to aircraft movement that a pilot is able to achieve in performing a task. (Pilot vehicle performance is a measure of handling performance. Pilot performance is a measure of the manner or efficiency with which a pilot moves the principal controls in performing a task.)

HANDLING QUALITIES
Those qualities or characteristics of an aircraft that govern the ease and precision with which a pilot is able to perform the tasks required in support of an aircraft role.

ROLE
The function or purpose that defines the primary use of an aircraft.

MISSION
The composite of pilot-vehicle functions that must be performed to fulfill operational requirements. May be specified for a role, complete flight, flight phase, or flight subphase.

TASK
The actual work assigned a pilot to be performed in completion of or as representative of a designated flight segment.

WORKLOAD
The integrated physical and mental effort required to perform a specified piloting task.
NASA Task Load Index (TLX)

- Subjective assessment on 6 scales
  - Mental demand
  - Physical demand
  - Temporal demand
  - Effort
  - Performance
  - Frustration level
- Rating from 1 (low) to 5 (high)

- 15 pairwise comparisons for assessment of relative importance of scales
- Final score is weighted average of scale values
### NASA Task Load Index (TLX)

#### Rating Scale Definitions

<table>
<thead>
<tr>
<th>Title</th>
<th>Descriptions</th>
</tr>
</thead>
<tbody>
<tr>
<td>MENTAL DEMAND</td>
<td>How much mental and perceptual activity was required (e.g., thinking, deciding, calculating, remembering, looking, searching, etc.)? Was the task easy or demanding, simple or complex, exacting or forgiving?</td>
</tr>
<tr>
<td>PHYSICAL DEMAND</td>
<td>How much physical activity was required (e.g., pushing, pulling, turning, controlling, activating, etc.)? Was the task easy or demanding, slow or brisk, slack or strenuous, restful or laborious?</td>
</tr>
<tr>
<td>TEMPORAL DEMAND</td>
<td>How much time pressure did you feel due to the rate or pace at which the tasks or task elements occurred? Was the pace slow and leisurely or rapid and frantic?</td>
</tr>
<tr>
<td>PERFORMANCE</td>
<td>How successful do you think you were in accomplishing the goals of the task set by the experimenter (or yourself)? How satisfied were you with your performance in accomplishing these goals?</td>
</tr>
<tr>
<td>EFFORT</td>
<td>How hard did you have to work (mentally and physically) in accomplishing your level of performance?</td>
</tr>
<tr>
<td>FRUSTRATION LEVEL</td>
<td>How insecure, discouraged, irritated, stressed and annoyed versus secure, gratified, content, relaxed and complacent did you feel during the task?</td>
</tr>
</tbody>
</table>

#### Scale Values

- **MENTAL DEMAND**
  - Low
  - High
- **PHYSICAL DEMAND**
  - Low
  - High
- **TEMPORAL DEMAND**
  - Low
  - High
- **PERFORMANCE**
  - Poor
  - Good
- **EFFORT**
  - Low
  - High
- **FRUSTRATION LEVEL**
  - Low
  - High
Four Designated Exploration Sites
EVA Walking Traverse
Approaching Science Site
Night Geology Exploration with Suit Lights
## Cooper-Harper Assessment Results

<table>
<thead>
<tr>
<th>Subject</th>
<th>Shirtsleeve</th>
<th>EVA/Walking</th>
<th>EVA/Rover</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>4</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>4</td>
<td>3</td>
</tr>
<tr>
<td>3</td>
<td>1</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>4</td>
<td>1</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>Average</td>
<td>1</td>
<td>3.5</td>
<td>4</td>
</tr>
<tr>
<td>Std. Dev.</td>
<td>0</td>
<td>0.58</td>
<td>0.82</td>
</tr>
</tbody>
</table>
NASA Task Load Index (TLX) Results
Team Techniques

- Behavioral Observation Scales
- Comms Usage Diagram
- Coordination Demands Analysis
- Team Decision Requirement Exercise
- Social Network Analysis
- Team Cognitive Task Analysis
- Team Communications Analysis
- Team Task Analysis
- Team Workload Assessment
- Task and Training Requirements Methodology
Team Workload Assessment

• Builds off of NASA TLX technique
• Team performs mission/run/sortie/simulation
• Individuals fill out NASA TLX
• Group fills out TLX as a team, from point of view of team
• Compare ratings between individuals, and between individuals and team
• Best indicator has been shown to be rating of poorest-performing individual
Interface Analysis Techniques

- Checklists
- Heuristics
- Interface Surveys
- Layout Analysis
- Link Analysis
- Questionnaire for User Interface Satisfaction
- Repertory Grids
- Software Usability Measurement Inventory
- System Usability Scale
- User Trials
- Walkthrough Analysis
Walkthrough Analysis

- Experienced system operators perform “walkthrough” of system under consideration
- Subject (pretends to) perform each required action, explaining function of each control and display used
- Analyst(s) can halt walkthrough and ask questions at any point
- Highly useful in early stages of development
- Highly useful for teams of operators
- Requires experienced operator(s)
System Design Techniques

- Allocation of Functions Analysis
- Focus Groups
- Groupware Task Analysis
- Mission Analysis
- Scenario Based Design
- Task Centered System Design
Scenario Based Design

- Create a series of operating scenarios for systems
- Use Walkthrough Analysis approach to “simulate” actions of experienced operators
- Determine actions, possible outcomes, design modifications to improve results
- Many potential scenarios need to be created and “played out”
- Only as good as experts and scenario designers
Performance Time Assessment Techniques

- Keystroke Level Method
- Timeline Analysis
- Critical Path Analysis
Timeline Analysis

- Collect operating timeline data wherever possible
- Perform HTA on system
- Link performance time estimates to bottom-level HTA goals
- Construct a timeline graph of nominal and off-nominal performance of mission
- Workload analysis maps directly into timelines
- Predictive performance depends on error-free execution
Pain Scales

• Human factors testing can cause distress, discomfort, and/or pain to test subject
• Various scales have been established to “quantify” pain levels
• Highly subjective, differs from person to person
Sample Pain Rating Scales

**UNIVERSAL PAIN ASSESSMENT TOOL**

This pain assessment tool is intended to help patient care providers assess pain according to individual patient needs. Explain and use 0-10 Scale for patient self-assessment. Use the faces or behavioral observations to interpret expressed pain when patient cannot communicate his/her pain intensity.

0 1 2 3 4 5 6 7 8 9 10

**Verbal Descriptor Scale**

0 = NO PAIN
1 = MILD PAIN
2 = MODERATE PAIN
3 = MODERATE PAIN
4 = SEVERE PAIN
5 = WORST PAIN POSSIBLE

**Wong-Baker Facial Grimace Scale**

0 = Alert Smiling
1 = No humor, serious flat
2 = Furrowed brow, pursed lips, breath holding
3 = Wrinkled nose, raised upper lips, rapid breathing
4 = Slow blink, open mouth
5 = Eyes closed, moaning, crying

**Activity Tolerance Scale**

0 = NO PAIN
1 = CAN BE IGNORED
2 = INTERFERES WITH TASKS
3 = INTERFERES WITH CONCENTRATION
4 = INTERFERES WITH BASIC NEEDS
5 = BEDREST REQUIRED
References

• Neville A. Stanton, P. M. Salmon, G. H. Walker, C. Baber, and D. P. Jenkins, Human Factors Methods: A Practical Guide for Engineering and Design - Ashgate, 2005

• Alphonse Chapanis, Human Factors in Systems Engineering - Wiley Interscience, 1996