Scenario Design: Designing Conditions Under Which Performance is Evaluated

Emilie M. Roth
Roth Cognitive Engineering
User-in-the-loop evaluations are key to insure that new support systems effectively aid performance.

Central to the design of meaningful evaluations is the ability to articulate the ‘*model of support*’:
- What situations is the system intended to support?
- What cognitive or collaborative challenges is it intended to address?

When you design evaluations it is important to *craft* evaluation scenarios that enable you to effectively test this model of support
Common Pitfalls in Scenario Design

- Use of ‘textbook’ (straightforward) use case scenarios.
- Over emphasis on ‘face validity’:
  - Actual data streams
  - Dynamic simulations

Problems:
- No guarantee that the sampled situations are representative of the cognitive and collaborative domain challenges
- No guarantee that the sampled situations are representative of the conditions where the DSS will be effective – reflect the model of support.
- Risk of getting floor or ceiling effects
HSI Approach to Scenario Design

- Requires designers to be explicit about the specific cognitive and collaborative activities to be supported.

- Enables pointed testing of the effectiveness of the proposed aiding concepts: *The Model of Support*
  - Are the hypothesized positive impacts realized?
  - What are additional unanticipated requirements for support?
Hypothetical Example: Decision Support System for Scheduling and Managing Multiple UAVs

- **Model of Support:**
  - User displays support immediate SA of task allocations and status across UAVs.
    - What each UAV is doing, what it is scheduled to do next, whether it’s current status is compatible with its current and future task assignments.
  - Automated planner enables users to dynamically re-allocate UAV assets to support original mission objectives with changed UAV status.
# Crafting Scenarios to test ‘Model of Support’

## 1. Situation Awareness of UAV status relative to plan

**Situation:** Deviations from expectations underlying mission plan

**Hypothesis of Support:**
User displays support immediate SA of task allocations and status across UAVs.

**Scenario Probe: Purpose - evaluate ability to assess a role/status mismatch**

A UAV has been assigned to surveil a given area, however, due to weather delays it’s fuel is insufficient to meet task objectives

**Decision Making Challenge:** The flight lead must recognize the mismatch

**Design Support Functions:** Observability of automated planning, observability of UAV current status
## 2. Dynamic Replanning

**Situation:** Deviations from expectations underlying mission plan

**Hypothesis of Support:** Automated planner enables users to dynamically re-allocate UAV assets

**Scenario Probe:** Purpose - exercise ability to adapt automated plan in response to plan/situation mismatch

The UAV cannot meet it’s surveillance objective, but there are other UAVs in the area that can take over.

**Decision Making Challenge:** The flight lead must recognize the mismatch and dynamically re-allocate tasks across the UAV assets by seeing available options.

**Design Support Functions:** Directability of automated planning
Multi-Facets to Scenario Design

Model of Support

- What cognitive and collaborative tasks is it intended to support?
- In what situations?

Complicating Situational Factors:

- Challenge situation assessment
- Challenge planning
- Challenge communication
- ...

Known cognitive and collaborative limitations:

- Decision-biases
- Memory & attention limits
- Change blindness
- ...

Inputs to Scenario Design
Fukushima Daiichi Accident

March 11, 2011  Tsunami height: approx 13 m

Loss of critical functions to prevent core damage and mitigate impacts
Created Complex Unanticipated Demands on Operators

- Total loss of sensor indications
- Dark, high radiation
- Available Procedures inapplicable
- Preplanned mitigation strategies inapplicable
- Limited ability to communicate between control room, Emergency Response Center, and operators in the field
- Need to assess, prioritize, and cope with problems in multiple units simultaneously
Sources of Input to Design of Scenarios

- Operational experience reviews
- Lessons from other industries
- Cognitive task analyses
- Actual cases provided by users
- Generic lists of complicating situational factors
We developed a list that can be used as ‘seeds’ for identifying domain complexities to incorporate in test scenarios (Patterson, Roth & Woods, 2010).

These factors are loosely organized around core Macrocognitive (cognitive and collaborative) functions:

- detection/monitoring;
- sense-making/situation assessment;
- planning/action formulation;
- information sharing/communication/coordination and attention/workload management.

Most recent embodiment of a working effort to capture characteristics that pose challenges to macrocognitive processes in a domain independent fashion.
### Example Complicating Situational Factors

<table>
<thead>
<tr>
<th>Macro-Cognitive Function</th>
<th>Complicating Situational Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Detecting / Noticing</td>
<td>Data overload</td>
</tr>
<tr>
<td></td>
<td>Missing Information</td>
</tr>
<tr>
<td>Diagnosing / ‘Sense-Making’</td>
<td>Ambiguous cues</td>
</tr>
<tr>
<td></td>
<td>‘Mismatch with expectations based on training / mental models</td>
</tr>
<tr>
<td>Planning/Deciding</td>
<td>Competing Goals</td>
</tr>
<tr>
<td></td>
<td>Mismatch with procedures</td>
</tr>
<tr>
<td>Communication / Coordination</td>
<td>Multiple competing demands</td>
</tr>
<tr>
<td></td>
<td>Mismatch with organizational structure</td>
</tr>
</tbody>
</table>
Benefits of HSI Approach to Scenario Design

- Allows design of scenarios that:
  - go beyond routine ‘textbook’ cases
  - sample realistically demanding situations that are vulnerable to known cognitive and collaborative processing limitations
  - provide pointed test of hypothesized support:
    - Allows objective assessment of the impact of the aid on performance in representative situations
    - Enables users to exercise the aid in representative situations allowing them to provide informed feedback.
Checklist in Evaluating Test Scenarios

- What is the model of support being tested? Do the validation scenarios provide opportunities to exercise and test this model of support?

- Do the scenarios capture the range of complicating factors that arise in the actual operational environment so as to assess extent and boundaries of effectiveness of support?

- Have the performance issues (e.g., potential vulnerabilities, biases, errors and breakdown points) that can impact the decisions and related cognitive and collaborative activities of interest been identified? Do the test scenarios create opportunities to assess the impact of the decision aid on these potential performance deficiencies?

- Have situations/probes/target events been embedded in the scenarios so as to create the opportunity for cognitive and collaborative activities of interest to be exercised in an observable manner?
Back-Up
# Levels of Scenario Design

<table>
<thead>
<tr>
<th>Level</th>
<th>Integrity Target</th>
<th>Indicators of Sufficient Integrity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Surface validity</td>
<td>Professionals representing the target user population are engaged in the scenario activities and consider them face valid</td>
<td>• Participants will play  &lt;br&gt;• Participants can describe scenario relevance to work  &lt;br&gt;• Participants recommend participation</td>
</tr>
<tr>
<td>Model of support</td>
<td>The primary justifications for the design of the software are explored with the scenario</td>
<td>• Range of performance across participants  &lt;br&gt;• Scenario includes <em>boundaries of support</em> – where the support system may breakdown.  &lt;br&gt;• Evaluators agree on capability gaps</td>
</tr>
<tr>
<td>Justification for implementation</td>
<td>The evaluation supports the ability to defend implementation decisions against factions who want to block implementation</td>
<td>• Performance improvements would reap significant benefits to the organization and people in field  &lt;br&gt;• Objective “quantitative” measures distinguish one configuration from alternatives</td>
</tr>
<tr>
<td>Representative complexity</td>
<td>The scenario set covers the range of anticipated complexity in the field setting, from nominal cases to challenging cases</td>
<td>• Scenarios sample range of complicating probes found in CTA studies  &lt;br&gt;• Range in performance across participants</td>
</tr>
<tr>
<td>Performance observability</td>
<td>The scenario includes situations/probes/target events that create the opportunity for cognitive and collaborative activities of interest to be exercised in an <em>observable</em> manner</td>
<td>• Externally observable utterances of actions allow reliably determining quality of cognitive (individual) and collaborative (team) performance</td>
</tr>
</tbody>
</table>
References

Summary

Two important goals in test design:

- Evaluate the ‘model of support’ (hypothesis) embodied in prototype/decision-aid.
- Discover additional domain demands and unanticipated requirements to propel further design innovation
- Test Scenarios should be designed to support these goals:
  - capture the rage of complicating factors that arise in the actual operational environment
  - create opportunities to assess the impact of decision-aid on potential performance vulnerabilities