

## **NDM8**

### **Evaluating the Resilience of a Human-Computer Decision-Making Team: A Methodology for Decision-Centered Testing**

Topic: Emerging methodological and theoretical issues, controversies, and challenges:  
Metrics for macrocognition

Scott S. Potter, William C. Elm, and James S. Tittle

Point of Contact: Scott S. Potter

Cognitive Systems Engineering Center  
ManTech Corporation

436 Seventh Ave., Suite 200

Pittsburgh, PA 15219

412-471-3456x20 (voice) 412-471-3461 (fax)

[Scott.Potter@ManTech.com](mailto:Scott.Potter@ManTech.com)

## **EXTENDED ABSTRACT**

In developing new decision support systems, it is essential to evaluate the impact of the new system on the decision-making it was designed to support. Embracing the perspective of macrocognition – the cognitive functions that are performed in natural decision-making settings (Cacciabue and Hollnagel, 1995) – imposes significant challenges for evaluation of new decision support systems. This perspective requires using measures that go well beyond typical performance metrics such as the number of subtasks achieved per person per unit of time and the corresponding simple baseline comparisons or workload assessment metrics. For instance, for a new military command post workstation, typical performance measures might be number of enemy tanks destroyed per period of watch. This procedure is likely to involve standard scenarios to facilitate comparisons between systems, but does not address the resilience of the JCS in the face of novel situations outside the boundaries. However, macrocognition-grounded evaluation frameworks, corresponding operational techniques, and corresponding metrics are limited.

In response to the need to evaluate the decision-making effectiveness of a practitioner teamed with a new decision support system, we have developed an evaluation methodology based on principles from Cognitive Systems Engineering (CSE). Decision Centered Testing (DCT; Rousseau, et al., 2005) aims at testing the effectiveness of operators teamed with Decision Support Systems (DSS) in any challenging work domain. DCT is grounded in a CSE framework, where the concept of a Joint Cognitive System (JCS) is central. We define a Joint Cognitive System (JCS) as the combination of human problem solver and automation/technologies which must act as co-agents to achieve goals and objectives in a complex work domain (cf. Hollnagel and Woods, 2005 for their cyclic model of the function of Joint Cognitive Systems). An implication of our Cognitive Systems Engineering foundation is that a JCS needs to be evaluated for its effectiveness in performing the complex cognitive work requirements. This JCS perspective implies that the system must be designed and evaluated from the perspective of the teaming of the practitioner and the DSS. Previous research in CSE and our own experience has lead us to identify a set of generic JCS support requirements that apply to cognitive work by any cognitive agent or any set of cognitive agents, including teams of people and machine agents. Metrics will have to reflect such phenomena as “resilience” of a JCS. This places new burdens on evaluation techniques and frameworks, since metrics should be generated from a principled approach and based on fundamental principles of interest to the designers of the JCS.

DCT involves explicit design and analysis of tests based on the key decision making demands within the naturalistic work domain. The result is an explicit test design describing the cognitive problem under test, the hypothesized “edge” or latent potential weaknesses in the JCS, as well as the events that need to be included in the scenario. In DCT, test scenarios are developed to specify a progressive evolution of events that would be expected to stress the defined edge within the JCS. This decision-centered approach to testing has proven effective in discovering fundamentally new ways for evaluating the net decision-making effectiveness of the joint human-technology decision-making team. Ultimately, DCT is about evaluating the resilience of the JCS.

The critical aspects of this technique are:

- The evaluation needs to address the attributes / characteristics that a joint cognitive system must have in order to effectively conduct cognitive work in any goal-directed task (Woods and Hollnagel, 2006). These include requirements such as observability, directability, teamwork with agents, and directed attention. For example, testing for observability might entail assessing the insights into a process that the user was able to gain through the representation provided by the DSS.
- The evaluation needs to be focused on the fundamental cognitive functions of the natural decision-making setting. To meet this need, we have found significant leverage from an analytical model of cognitive demands of the work domain derived from a cognitive work analysis. This has provided a pragmatic approach to focus the evaluation on the most profitable functions (since an exhaustive evaluation is infeasible).
- In order to evaluate the resilience of the JCS, it is important to define potential weaknesses, or ‘fissures’ in the JCS where decision-making effectiveness may break down under pressure. Within the DCT methodology, these weaknesses are referred to as ‘edges’. An edge is any discontinuity within any of the relationships within a JCS.
- Once the target of the DCT has been identified, scenarios can be defined explicitly on the analytical basis in order to exercise the desired cognitive demands. Scenarios can be derived from the specific ‘edge’ under investigation. For example, a DCT focused on selection between competing goals would include scenarios specifically designed to emphasize the different goals.
- “Cognitive pressure” can be defined to explicitly stress the edges and therefore assess the resilience of the JCS. Cognitive pressure is not simply a ‘swamping’ or overloading the JCS by the number of stimuli or amount of cognitive work to be performed. Rather, it is a carefully crafted manipulation of controllable aspects of the work environment related to JCS support requirements, cognitive demands of the work domain, and/or edges in the JCS.

An initial application of the DCT methodology enabled the definition of appropriate test metrics and the construction of unique test scenarios to exercise the decision-making effectiveness. From this application, significant insights were gained regarding the construction of an evaluation framework for assessing the net decision-making effectiveness of a C2 system.

## REFERENCES

- Cacciabue, P. C. and Hollnagel, E. (1995). Simulation of cognition: Applications. *Expertise and Technology: Cognition and Human-Computer Cognition*. J. M. Hoc, P. C. Cacciabue, and E. Hollnagel, eds. Lawrence Erlbaum Associates.
- Hollnagel, E. and Woods, D. D. (2005). *Joint cognitive systems: Foundations of cognitive systems engineering*. Boca Raton, FL: Taylor and Francis.
- Rousseau, R., Easter, J., Elm, W., and Potter, S. (2005). Decision-Centered Testing (DCT): Evaluating Joint Computer Cognitive Work. In *Proceedings of the Human Factors and Ergonomics Society 49th Annual Meeting*. Santa Monica, CA: HFES.
- Woods, D. D. and Hollnagel, E. (2006). *Joint cognitive systems: Patterns in cognitive systems engineering*. Boca Raton, FL: Taylor and Francis.