14TH ICCRTS

C2 AND AGILITY

Supporting the C2 of Training Exercise Management: The design of a Decision-Centered Scenario Management Tool

Topics: (8) C2 Assessment Tools and Metrics

Scott S. Potter and Michael P. Farry

Point of Contact: Scott S. Potter, PhD

Charles River Analytics
625 Mount Auburn St.
Cambridge, MA 02138
617-491-3474x648 (voice) 866-571-2631 (fax)
spotter@cra.com

ABSTRACT

In this effort, we investigated the problem of supporting the real-time management of complex scenarios within a large-scale training or evaluation exercise. Developing cognitively challenging scenarios is a fundamental challenge for these types of situations (Mumaw and Roth, 1992; Roth, et al., 2002). For the exercise or evaluation to be useful and effective, the scenarios must be carefully designed to address specific training objectives as well as specific cognition and collaboration demands imposed on the participants. Then, the scenario must be able to be managed efficiently during the exercise to adapt to disruptions and deviations from the plan. The critical issues identified through our investigation and addressed in our proof-of-concept demonstration include:

- *Designing and managing cognitively challenging scenarios.* Critical to the effective exercise of cognitive skills is the design and conduct of challenging scenarios. This includes specification of the (often) complex set of world events to set up the challenging decision-making situation as well as the modification of these events as the training exercise unfolds.
- *Representing System Events and Activity.* Fully-specified scenarios require the coordination of system activity in response to world events. This includes automated system activity as well as information processing activity in support of human decision-making.
- *Depicting Operator Activity.* Both predicted and actual activity in response to world events and system activity should be represented. Specifying expected behavior provides the ability to identify deviations from planned behavior as well as confirmation that the progression of operator activity is going as planned.
- *Linking Scenario Events and Activity to Exercise Objectives.* This is critical for the determination of achieving training objectives. Establishing these links supports the ability to determine the significance of world events in exercising the training objectives. In addition, this linkage permits the planning and management of the training exercise to meet the desired objectives.

In this paper, we discuss these concepts and directions for potential follow-on efforts to implement and evaluate these concepts. Note that these concepts should apply widely and generally in the design and evaluation of software, work systems, and training systems.

Introduction

Training warfighters in large-scale exercises is essential to maintain their readiness at an acceptable level. In particular, simulated training exercises can provide invaluable experience for the warfighters and significantly mitigate the cost and risk of live, force-on-force training. Millions of dollars have been spent over the past several years on developing advanced training systems, training infrastructures, simulations, and Live, Virtual, and Constructive (LVC) training system interoperability. However, the success of a training exercise still depends on the ability of training managers to easily and rapidly digest the diverse, complex, and interacting streams of information that is inherent to large-scale, LVC distributed simulation-based training exercises and their ability to craft and conduct well-orchestrated scenarios within the training system.

For example, a large-scale training exercise might include hours of events and activities leading to a critical, culminating event (e.g., simulated forces involved in a deception maneuver). If any of the key entities in the culminating event do not behave as expected (e.g., take longer to complete their tasks than anticipated, fail to complete their tasks), the culminating event may not be able to proceed as planned, and all of the build-up time will have been wasted. Unfortunately, the scenario deviations are often not noticed until too late because there is no effective system that is managing the streams of information to effectively deal with unanticipated events – there is no “knowledge-optimized Common Operating Picture” (COP) specific to management of the training exercise scenario. As a result, the instructional value of the exercises depends on the training manager’s ability to rapidly understand information about the entities and coordinate these entities’ behavior to produce a well-orchestrated training exercise.

In planning and conducting these training exercises, a typical approach is to develop training objectives and then create the specific scenarios and instantiating these scenarios in a Master Scenario Event List (MSEL). Unfortunately, the MSEL typically lacks any links to the underlying training objectives (only contains references) as well as any links to the decision-making behavior being exercised by the scenario. In addition, there is no representation of the activity of the command and control (C2) system in responding to the events in the scenario (a critical aspect of evaluating a Joint Cognitive System – human problem solvers teamed with automation technologies in a complex work domain such as C2 (cf. Hollnagel and Woods, 2005)).

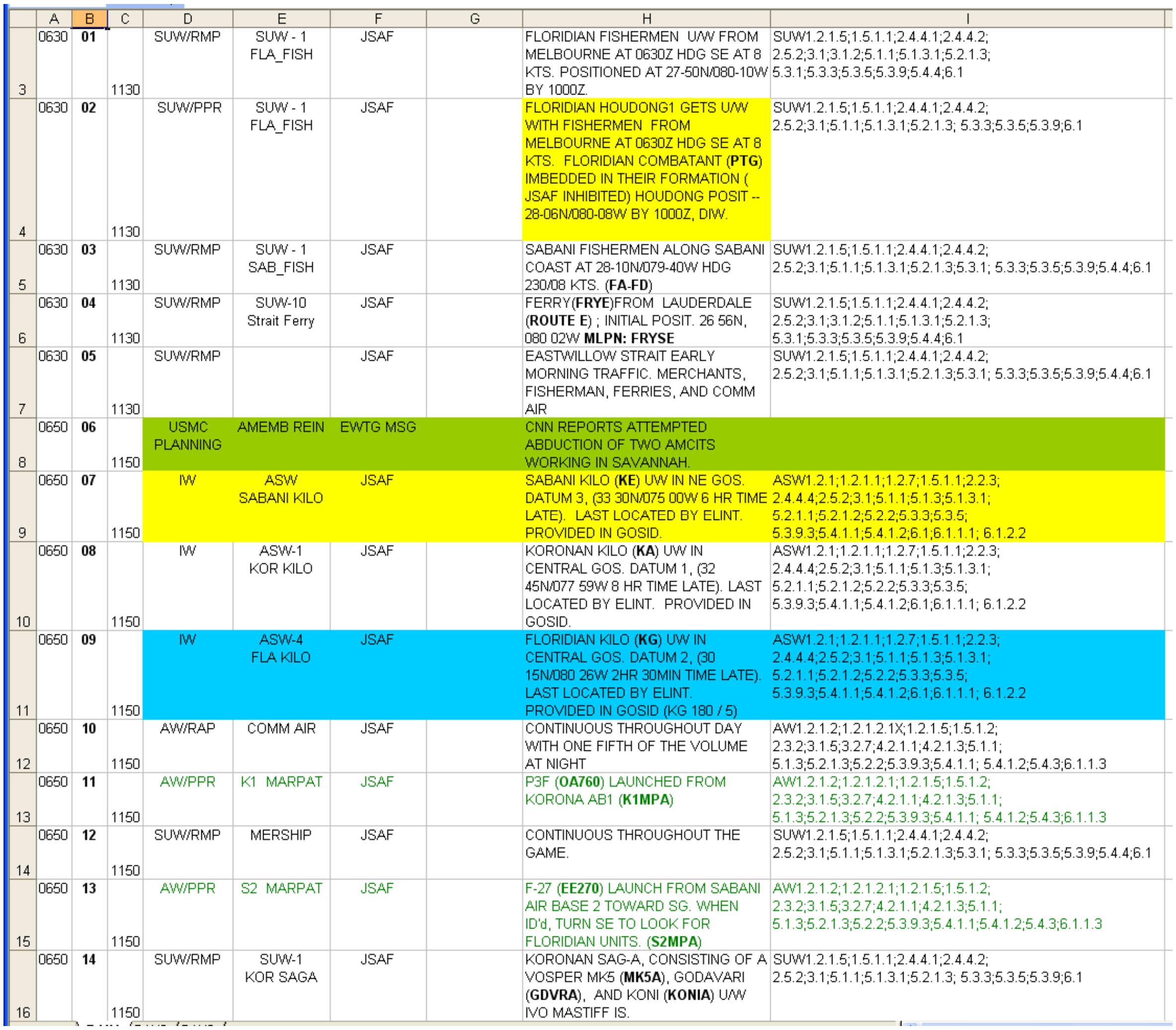
As indicated in Figure 1 (a representative section from a MSEL), the MSEL is a very compact representation of the scenario events, and therefore an under-specification of the critical issues related to designing effective scenarios, representing the essential characteristics of scenarios, and managing these scenarios in real-time. Some specific weaknesses include:

- All events are represented as instantaneous, based on a timestamp. This is an under-specification and misrepresentation of the temporal duration of events. In real-world situations, “events” sometimes are defined by the fact that they do not have clear-cut (or even knowable) beginnings or end points.
- Related events (multiple events from the same entity) are not represented except for repeating the ID of the entity. Real-world situations are always wrapped

inside other situations and contexts, and typically inside multiple contexts and larger-scale situations.

- Exercise objectives addressed by an event are indicated by textual links that make identifying specific objectives very difficult.
- Important events are highlighted in several ways (background and text color) without any indication of the meaning of the highlighting. This is always a tipping point for breakdown in operator effectiveness.
- Temporal relationships between events are not indicated. This could include temporal dependencies and proximity. Our technology and means of training generally presume that events and activities can be temporally bounded in clear and often precise ways. The real world typically does not work that way. While scenario and training design procedures and technologies presume the specification of timelines, those need not necessarily translate into an infidelity in the actual unfolding of the training or scenario events.
- There is no distinction between world events and system events (in response to the world events) and no representation of anticipated responses to the events.
- There is no indication for the trainer of the rationale behind the events, conveying the reason for including the specific event(s) or the intended decisions or behavior being induced. This is a critical point. Often in design, and design documentation, the rationale for design elements is lost in the engineering descriptions. This *always* creates issues and problems downstream.
- There is no provision for deviations from the MSEL based on participant behavior (e.g., delayed response, unexpected behavior, communications link failures, hang-up of software modules, spontaneously corrupted data files, etc, multiple plausible responses).
- There is no ability to adapt the scenario in real-time based on the progression of events or have multiple potential paths through the scenario. This too represents a tipping point for lack of fidelity.

In general, there is an irony. While training scenarios are intended to bring the complexity of the world into the training situation, the means, procedures, and technological constraints for the design and conduct of training can result in training experiences that are misrepresentative of the complexity of the real-world, sometimes dangerously misrepresentative and therefore actually counter-productive.



	A	B	C	D	E	F	G	H	I
3	0630	**01**	1130	SUW/RMP	SUW - 1 FLA_FISH	JSAF		FLORIDIAN FISHERMEN U/W FROM MELBOURNE AT 0630Z HDG SE AT 8 KTS. POSITIONED AT 27-50N/080-10W BY 1000Z.	SUW1.2.1.5;1.5.1.1;2.4.4.1;2.4.4.2; 2.5.2;3.1;3.1.2;5.1.1;5.1.3.1;5.2.1.3; 5.3.1;5.3.3;5.3.5;5.3.9;5.4.4;6.1
4	0630	**02**	1130	SUW/PPR	SUW - 1 FLA_FISH	JSAF		FLORIDIAN HOUDONG1 GETS U/W WITH FISHERMEN FROM MELBOURNE AT 0630Z HDG SE AT 8 KTS. FLORIDIAN COMBATANT (**PTG**) IMBEDDED IN THEIR FORMATION (JSAF INHIBITED) HOUDONG POSIT -- 28-06N/080-08W BY 1000Z, DIW.	SUW1.2.1.5;1.5.1.1;2.4.4.1;2.4.4.2; 2.5.2;3.1;5.1.1;5.1.3.1;5.2.1.3; 5.3.3;5.3.5;5.3.9;6.1
5	0630	**03**	1130	SUW/RMP	SUW - 1 SAB_FISH	JSAF		SABANI FISHERMEN ALONG SABANI COAST AT 28-10N/079-40W HDG 230/08 KTS. (**FA-FD**)	SUW1.2.1.5;1.5.1.1;2.4.4.1;2.4.4.2; 2.5.2;3.1;5.1.1;5.1.3.1;5.2.1.3;5.3.1; 5.3.3;5.3.5;5.3.9;5.4.4;6.1
6	0630	**04**	1130	SUW/RMP	SUW-10 Strait Ferry	JSAF		FERRY(**FRYE**)FROM LAUDERDALE (**ROUTE E**) ; INITIAL POSIT. 26 56N, 080 02W **MLPN: FRYSE**	SUW1.2.1.5;1.5.1.1;2.4.4.1;2.4.4.2; 2.5.2;3.1;3.1.2;5.1.1;5.1.3.1;5.2.1.3; 5.3.1;5.3.3;5.3.5;5.3.9;5.4.4;6.1
7	0630	**05**	1130	SUW/RMP		JSAF		EASTWILLOW STRAIT EARLY MORNING TRAFFIC. MERCHANTS, FISHERMAN, FERRIES, AND COMM AIR	SUW1.2.1.5;1.5.1.1;2.4.4.1;2.4.4.2; 2.5.2;3.1;5.1.1;5.1.3.1;5.2.1.3;5.3.1; 5.3.3;5.3.5;5.3.9;5.4.4;6.1
8	0650	**06**	1150	USMC PLANNING	AMEMB REIN	EWTG MSG		CNN REPORTS ATTEMPTED ABDUCTION OF TWO AMCITS WORKING IN SAVANNAH.	
9	0650	**07**	1150	IW	ASW SABANI KILO	JSAF		SABANI KILO (**KE**) UW IN NE GOS. DATUM 3, (33 30N/075 00W 6 HR TIME LATE). LAST LOCATED BY ELINT. PROVIDED IN GOSID.	ASW1.2.1;1.2.1.1;1.2.7;1.5.1.1;2.2.3; 2.4.4.4;2.5.2;3.1;5.1.1;5.1.3;5.1.3.1; 5.2.1.1;5.2.1.2;5.2.2;5.3.3;5.3.5; 5.3.9.3;5.4.1.1;5.4.1.2;6.1;6.1.1.1; 6.1.2.2
10	0650	**08**	1150	IW	ASW-1 KOR KILO	JSAF		KORONAN KILO (**KA**) UW IN CENTRAL GOS. DATUM 1, (32 45N/077 59W 8 HR TIME LATE). LAST LOCATED BY ELINT. PROVIDED IN GOSID.	ASW1.2.1;1.2.1.1;1.2.7;1.5.1.1;2.2.3; 2.4.4.4;2.5.2;3.1;5.1.1;5.1.3;5.1.3.1; 5.2.1.1;5.2.1.2;5.2.2;5.3.3;5.3.5; 5.3.9.3;5.4.1.1;5.4.1.2;6.1;6.1.1.1; 6.1.2.2
11	0650	**09**	1150	IW	ASW-4 FLA KILO	JSAF		FLORIDIAN KILO (**KG**) UW IN CENTRAL GOS. DATUM 2, (30 15N/080 26W 2HR 30MIN TIME LATE). LAST LOCATED BY ELINT. PROVIDED IN GOSID (KG 180 / 5)	ASW1.2.1;1.2.1.1;1.2.7;1.5.1.1;2.2.3; 2.4.4.4;2.5.2;3.1;5.1.1;5.1.3;5.1.3.1; 5.2.1.1;5.2.1.2;5.2.2;5.3.3;5.3.5; 5.3.9.3;5.4.1.1;5.4.1.2;6.1;6.1.1.1; 6.1.2.2
12	0650	**10**	1150	AW/RAP	COMM AIR	JSAF		CONTINUOUS THROUGHOUT DAY WITH ONE FIFTH OF THE VOLUME AT NIGHT	AW1.2.1.2;1.2.1.2.1X;1.2.1.5;1.5.1.2; 2.3.2;3.1.5;3.2.7;4.2.1.1;4.2.1.3;5.1.1; 5.1.3;5.2.1.3;5.2.2;5.3.9.3;5.4.1.1; 5.4.1.2;5.4.3;6.1.1.3
13	0650	**11**	1150	AW/PPR	K1 MARPAT	JSAF		P3F (**OA760**) LAUNCHED FROM KORONA AB1 (**K1MPA**)	AW1.2.1.2;1.2.1.2.1;1.2.1.5;1.5.1.2; 2.3.2;3.1.5;3.2.7;4.2.1.1;4.2.1.3;5.1.1; 5.1.3;5.2.1.3;5.2.2;5.3.9.3;5.4.1.1; 5.4.1.2;5.4.3;6.1.1.3
14	0650	**12**	1150	SUW/RMP	MERSHIP	JSAF		CONTINUOUS THROUGHOUT THE GAME.	SUW1.2.1.5;1.5.1.1;2.4.4.1;2.4.4.2; 2.5.2;3.1;5.1.1;5.1.3.1;5.2.1.3;5.3.1; 5.3.3;5.3.5;5.3.9;5.4.4;6.1
15	0650	**13**	1150	AW/PPR	S2 MARPAT	JSAF		F-27 (**EE270**) LAUNCH FROM SABANI AIR BASE 2 TOWARD SG. WHEN ID'd, TURN SE TO LOOK FOR FLORIDIAN UNITS. (**S2MPA**)	AW1.2.1.2;1.2.1.2.1;1.2.1.5;1.5.1.2; 2.3.2;3.1.5;3.2.7;4.2.1.1;4.2.1.3;5.1.1; 5.1.3;5.2.1.3;5.2.2;5.3.9.3;5.4.1.1;5.4.1.2;5.4.3;6.1.1.3
16	0650	**14**	1150	SUW/RMP	SUW-1 KOR SAGA	JSAF		KORONAN SAG-A, CONSISTING OF A VOSPER MK5 (**MK5A**), GODAVARI (**GDVRA**), AND KONI (**KONIA**) U/W IVO MASTIFF IS.	SUW1.2.1.5;1.5.1.1;2.4.4.1;2.4.4.2; 2.5.2;3.1;5.1.1;5.1.3.1;5.2.1.3; 5.3.3;5.3.5;5.3.9;6.1

Figure 1. Representative Master Scenario Event List (MSEL)

A better, as well as innovative decision support environment is clearly needed to improve the management of these training exercises. We see three distinct requirements for an effective solution. *First*, an informative and integrated “Exercise Common Operating Picture” is needed to represent the various entities, systems, and participants in the training exercises because the training managers must have a high degree of situation awareness about the behavior of these entities to assess whether they are performing as expected and if the training objectives are being met. *Second*, better mechanisms are needed to control the entities in the training exercises because training managers must insert, delete, redirect, and modify the entities and their behavior to ensure they are meeting the training objectives. *Third*, intelligent data management technology is needed to monitor the entities in the training exercise because the existing entities do not communicate information about their behavior or performance. As part of this requirement, technology needs to capture, integrate, and disseminate information about entity performance in the context of the training objectives, and convey this in a meaningful way to the managers of the training exercises It must allow training managers to immediately and directly perceive when entities are not behaving as required, immediately apprehend the impact on training objectives, and rapidly re-direct other entities to take their place without any disruption to the training exercise.

1. Technical Approach and Results

1.1 Knowledge Elicitation

As a mechanism for bounding the problem and identifying critical insights related to the problem, we conducted knowledge elicitation interviews with Navy experimentation personnel. Based on these interviews, it became clear that the problems of managing training exercises are exacerbated by several practical and immediate factors.

1. Exercise planning and management.

First, there is an apparent lack of rigor in the planning process contrasted with a wealth of experience on the part of the exercise planners in creating scenarios to address the specific training objectives. Given this combination, the planners are able, based on their vast experience, to quickly construct the training scenarios by adapting previously-used scenarios to fit the current objectives. Unfortunately, this can result in weak or broken mappings between scenario events and training objectives.

Second, it is not always the same team controlling the exercise as who planned the event. Also, the exercise management team can be different from the planning team. This can cause any implicit links between scenario events and training objectives or design rationale for the scenario can be lost in the transition between teams. Training managers can get confused and ask themselves, “why did they do it that way?” “What happens next? as the exercise begins to unravel.

Third, the training objectives are often ill-defined as well as being over-constrained. Too often, the exercises are defined in a bottom-up process, resulting in under-specified objectives. In addition, there are often multiple competing constraints on the training exercise. For example, new technology introduced into the experimentation environment can impose constraints on the training objectives. In addition, there are constraints imposed by the operational resources available for the exercise (e.g., number and type of live vs. simulated entities). Therefore, the objectives are determined by trying to best satisfy the constraints of the exercise environment, the new technology to be evaluated, and the operational resources available.

Fourth, there is a general difficulty in establishing the links between scenario events and the exercise objectives. This stems from a lack of understanding about what to evaluate (i.e., how to construct scenarios that exercise the desired objectives) and how to evaluate (i.e., what metrics to use to assess the performance being exercised by the scenario).

2. Scenario design and management.

First, there is no support or general guidance for designing effective scenarios; either for crafting them around the desired exercise objectives (as discussed above) or designing them in a coherent manner (e.g., to ensure they have the desired phases and transitions).

It is understood that scenarios need to have an underlying design basis, but typically they get generated based on prior use and experience.

Second, there is no support for modifying the scenario based on real-time unanticipated deviations from the original plan. These deviations could include changes in resources available (e.g., resource not available as expected; substitute resource required), participant(s) performance within the scenario (e.g., taking longer to complete activities; delay in next set of events required), or unexpected events occurring within the scenario (e.g., entity deviating from scripted behavior).

Third, scenarios focus on the system events (e.g., ISR task request generated) rather than the context and the precipitating events in the world that triggered these events. In addition, scenarios do not include any indication of the planned / actual participant behavior, despite the certain construction of expected responses to the scenario events by the exercise planners. Having these three components of the scenario was seen as valuable but not available in any of the current tools.

Fourth, scenarios do not pay sufficient attention to the decision-making demands imposed on the warfighters by the system events. As a result, there is a need for support for managing the relationships between events and decision-making demands. In addition, there is a need for integrating decision-centered measures and metrics into the scenarios (as the measures and metrics may impact the types of events to be used in the scenarios).

Critical to the effective exercise of cognitive skills is the design and conduct of challenging scenarios. This includes specification of the (often) complex set of world events to set up the challenging decision-making situation as well as the modification of these events as the training exercise unfolds. In the next section we describe our effort to define support requirements for designing effective scenarios.

1.2 Designing Effective Scenarios

In order to define requirements for our design concept, we synthesized related work on scenario design to identify critical factors in designing effective scenarios from the perspective of challenging the decision-making of the participants. From this effort we have defined the following premises:

1. Scenarios should be designed to exercise the Joint Cognitive System.

In training as well as evaluation, it is essential to assess the impact of new technology on the decision-making it was designed to support. Embracing the perspective of macrocognition, defined in complementary ways as (1) the cognitive functions that are performed in natural decision-making settings (Cacciabue and Hollnagel, 1995) or (2) how cognition adapts to complexity (Klein, 2007; Rasmussen and Lind, 1981) imposes significant challenges for scenario design and metric development. This perspective requires using measures that go 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 naval command and control system exercise, typical performance measures might be the percent of correct identifications and proximity of engagement to optimal probability of kill within the period of watch. This approach is likely to involve standard scenarios to facilitate comparisons between systems, but does not include any way to measure the resilience of the joint cognitive system (JCS; defined 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 a JCS) in the face of novel situations outside the typical performance boundaries. However, in order to impact these types of exercises, macrocognition-grounded frameworks must provide guidance with respect to defining the scenario characteristics and corresponding metrics. Macrocognition is defined in complementary ways as (1) the cognitive functions that are performed in natural decision-making settings (Cacciabue and Hollnagel, 1995) or (2) how cognition adapts to complexity (Klein, 2007; Rasmussen and Lind, 1981). This JCS perspective implies that the exercise must be designed from the perspective of the teaming of the practitioner and any supporting technologies.

Previous research in Cognitive Systems Engineering (CSE) and our own experience has led us to identify a set of support requirements that apply to cognitive work by any cognitive agent or any set of cognitive agents, including teams of people and machine agents (Billings and Woods, 1994; Dekker and Woods, 1999: Christoffersen and Woods, 2002; Elm, et al., 2005). These include:

- **Observability –** the ability to form insights into a process (either a process in the work domain or in the automation), based on feedback received. Observability overcomes the ‘keyhole’ effect and allows the practitioner to see sequences and evolution over time, future activities and contingencies, and the patterns and relationships in a process.
- **Directability –** the ability to direct/redirect resources, activities, and priorities as situations change and escalate. Directability allows the practitioner to effectively control the processes in response to (or in anticipation of) changes in the environment.
- **Teamwork with agents –** the ability to coordinate and synchronize activity across agents (both computer and human agents). This defines the type of coordination (e.g., seeding, reminding, critiquing) between agents. Teamwork with agents allows the practitioner to effectively re-direct agent resources as situations change.
- **Directed attention –** the ability to re-orient focus in a changing world. This includes issues like tracking others’ focus of attention and their interruptability. Directed attention allows the human-system team to work in a coordinated manner, resulting in increased effectiveness.
- **Resilience –** the ability to anticipate and adapt to surprise and error. This includes issues such as failure-sensitive strategies, exploring outside the current boundaries or priorities, overcoming the brittleness of automation, and maintaining peripheral awareness to maintain flexibility.

This places new burdens on scenario design, since scenarios and associated metrics should be generated from a principled approach and based on fundamental principles of interest to the designers of the JCS.

2. Scenarios should be designed to address practitioners' knowledge and reasoning skills.

The design of scenarios should be conducted using an understanding of individual and team knowledge and reasoning skills to assess decision making performance. Humans are capable of complex information processing; however, known cognitive biases can hamper the decision making process. For example, in information overload situations, decision-makers sometimes adopt mental strategies that simplify the problem, but which can lead to erroneous representations (Feltovich, Hoffman, Woods, and Roesler, 2004; Patterson, Roth, & Woods, 2001). The introduction of any technology into the decision-making process will have an impact – in some cases enhancing performance, and in others exacerbating performance problems. These knowledge and reasoning skills are important to consider when developing scenarios, because inclusion of event sequences meant to elicit these factors provide a more stringent test of the decision-making abilities of the participants.

Table 1 lists a sample of the types of individual and team knowledge and reasoning biases that can affect decision quality, as well as ways technology can mitigate or exacerbate the impact. Consideration of these biases can lead to testable hypotheses about the potential impact, both positive and negative, of new technologies, procedures, and training on the decision-making skills.

Table 1. Knowledge and Reasoning Biases

Knowledge and Reasoning Bias	Description	How Technology Can Mitigate Bias	How Technology Can Exacerbate Bias
Availability Bias (Tversky & Kahneman, 1973)	The ease with which humans can recall specific incidents affects judgments of frequency	Utilize context sensitive cues to determine the likelihood of an incident to aid in categorization	Present the incident history which will reinforce the likelihood that prior incidents will be more heavily weighted when making a decision
Primacy Effect (Anderson, 1981)	An order effect when information presented first is weighted more heavily than information presented later	Manipulate salience to ensure equal visibility of all information being utilized to make a decision	List information in the order which it was received, keeping the first piece of information at the top of the list
Recency Effect (Anderson, 1981)	The presentation of information just prior to making a decision overly influences the decision process	Manipulate salience to ensure equal visibility of all information being utilized to make a decision	Only display the most recent piece of information, hiding prior cues from view
Confirmation Bias (Watson, 1960)	Decision-makers tend to seek information consistent with their own opinion and ignore disconfirming evidence	Monitoring agents that alert user when the value for a particular cue falls outside of a pre-specified range	System architectures that allow for user to "drill-down" into the data, but no facilities for lateral movement
Frequency Bias (Eihorn & Hogarth, 1978)	Decision-makers tend judge the strength of the predictive relationship between two variables by focusing on the absolute frequency of events rather than their relative frequency (related to availability bias)	Dynamic process instrumented to represent function relationship between variables rather than the value of individual variables	Displays that provide informational values out of context
Concrete Information	Information that is vivid or	Displays that provide context	Self-organizing systems that morph

Bias (Nisbett & Ross, 1980)	based on personal experience tends to dominate abstract information like summaries or statistical base rates	dependent cue to help focus attention and alert user to unexpected or inconsistent findings	to match the user's model of the domain
Completeness Error (Russo, 1977)	Data displays that seem overly logical and complete that important omissions are overlooked	Displays that explicitly portray uncertainty to the user	High density interface
Conservatism (Edwards, 1968)	People often underestimate the value of new information	Utilize salience mapping to draw attention to those features which are most important within the problem space	System that allows the user to make decision prior to receiving all available information
Anchoring and Adjustment (Kahneman & Tversky, 1973)	Use initial information to anchor judgment and then adjust judgment as new information becomes available	Displays designed to represent cue salience to user	List information in the order which it was received
Fixation Errors/Failures to Revise Situation Assessment as Conditions Evolve Woods, et al., 2001)	Decision-makers fails to revise situation assessment as new evidence comes in	Critiquing Systems that flag evidence inconsistent with situation assessment	Advisory systems that are subject to the same biases as the decision-maker
Law of Small Numbers Tversky & Kahneman, 1971)	Decision makers tend to be overestimate the reliability of relatively small amounts of data	Displays that explicitly portray uncertainty to the user	Graphically dense displays that fail to make salient areas in which information is sparse
Question Format (Hogarth, 1975)	Estimates of probabilities have been found to differ according to the method with which they were asked to respond	Non-evaluative input tables that are based on a functional mapping of the domain	Context insensitive data entry forms
Scale Effects (Hogarth, 1975)	The scale on which responses are recorded can effect answers	Input formats based on a functional description of the information space	Context insensitive data entry forms
Fundamental Attribution Error (Nisbett & Ross, 1980)	People tend to attribute success to their own skill and failure to chance or the environment	Displays provide context sensitive visualizations that highlight and label the results of prior decisions	Utilizing god's-eye-views to display own actions and point of view displays for the actions of others
Hindsight Bias (Fischhoff, 1975)	After an event has occurred people will claim that predicted the event, even if prior to the event they were uncertain	Instrumented information suites that record operator communications and can be used to review team performance after decisions have been made	System that allow operators to comment on performance for some period after the decision has been made

3. Scenarios should be designed to exercise Complicating Factors

A second issue to consider when developing scenarios for training or evaluation is the inclusion of a range of "complicating factors." Scenarios should include not only routine or textbook cases, but also cases that challenge both individual cognition, as well as collaborative processes and reflect the real complexities that may arise in the domain. Complicating factors can be thought of as the types of conditions that can arise in real world situations (especially situations that the military encounters) to create opportunities for learning and adaptation.

There is a growing body of research that has attempted to capture and catalogue the types of complicating factors that arise in dynamic high risk domains which can impact the decision-making and collaborative processes of domain practitioners (Mumaw & Roth, 1992; Woods, Johannesen, Cook & Sarter, 1994; Roth, Mumaw & Lewis, 1994; Woods & Patterson, 2001). Complicating factors can be found across domains and provide challenging decision making and collaborative demands, which in turn provides a principled way to generate test scenarios that systematically probe potential decision-making vulnerabilities. Table 2 provides a list of typical complicating factors that arise in dynamic, high-risk work domains.

Table 2. Complicating Factors for Creating Cognitively Challenging Scenarios

Complicating Factors	Description / Scenario Characteristics
Garden path problems	Conditions start out with the situation appearing to be a simple problem (based on strong but incorrect evidence) and domain practitioners react accordingly. However, later correct symptoms appear, which the domain practitioners may not notice until it is too late.
Missing information	Key indicators may be missing due to failed sensors, lack of sensors, poor communication or lack of informants on the ground. For example may not know the location of friendly forces.
Misleading information	Misleading information may be provided due to inherent limitations of reports (e.g., stale information, inherent limitations of predictions, distortions resulting from indirect reports, secondary sources, translations) or explicit intent to deceive through misinformation. It can also result reliance on indirect indicators that are usually correlated with the information of interest, but not in that situation.
Masking Activities	Activities of other agents, or other automated systems may cover up or explain away key evidence.
Ambiguous Situations / Multiple lines of reasoning	Situations can occur where it is possible to think of significantly different explanations (e.g., in ambiguous situations) or response strategies, all of which seem valid at the time, but which may be in conflict (or a source of debate and disagreement by the operating crew).
Situations that change, requiring revised situation assessments	Once domain practitioners have developed a situation assessment and have started acting on it, it is often very difficult for them to recognize that there is new information or new conditions that requires them to change their situation assessment.
Side effects	Situations can arise where the effects of human or automated system actions, or effects of the initial failure, have side effects, which are not expected or understood.
Late changes in the plan	The situation is being managed according to a prepared plan, and then for some reason changes are required late in the situation. Domain practitioners can become confused as to next steps; the plan is no longer well tested and can contain flaws, or the whole "big picture" gets lost by those managing the event.
Impasses	The situation contains features where, at some point, it is very difficult for the domain practitioners to move forward, such as when the COA no longer matches the conditions, or assumed available personnel or resources are not available.
Trade offs	Domain practitioners must make impromptu judgments about choices between alternatives, such as when to wait to see if a problem develops (and may get out of control) versus jumping in early before it is clear what has caused the problem (just one of many examples).

Scenarios designed to address Complicating Factors can also address one or more of the Knowledge and Reasoning Biases, resulting in a rich specification of the design basis for the scenario. In this way, the design basis includes a dimension related to human decision-making characteristics (biases) and observable behaviors of the work domain (the Complicating Factors). This combination provides a rich skeletal structure for scenarios that is then instantiated with mission specific details that arise in the domain to produce a rich context for exercising decision-making effectiveness.

1.3 Decision Support Concept

In this effort, we have designed a scenario management tool to maintain these links in order to create a more effective means of developing, modifying, and exercising the scenario. In addition, we have developed a tool for establishing explicit links between the scenario events and the decision making / cognitive work requirements imposed by the events. The result is the ability to define decision-centered metrics for assessing performance within the scenario and maintain awareness of the training objectives being addressed by the particular phase of the scenario.

1. Design Concept

Our initial Scenario Management design concept is shown in Figure 2.

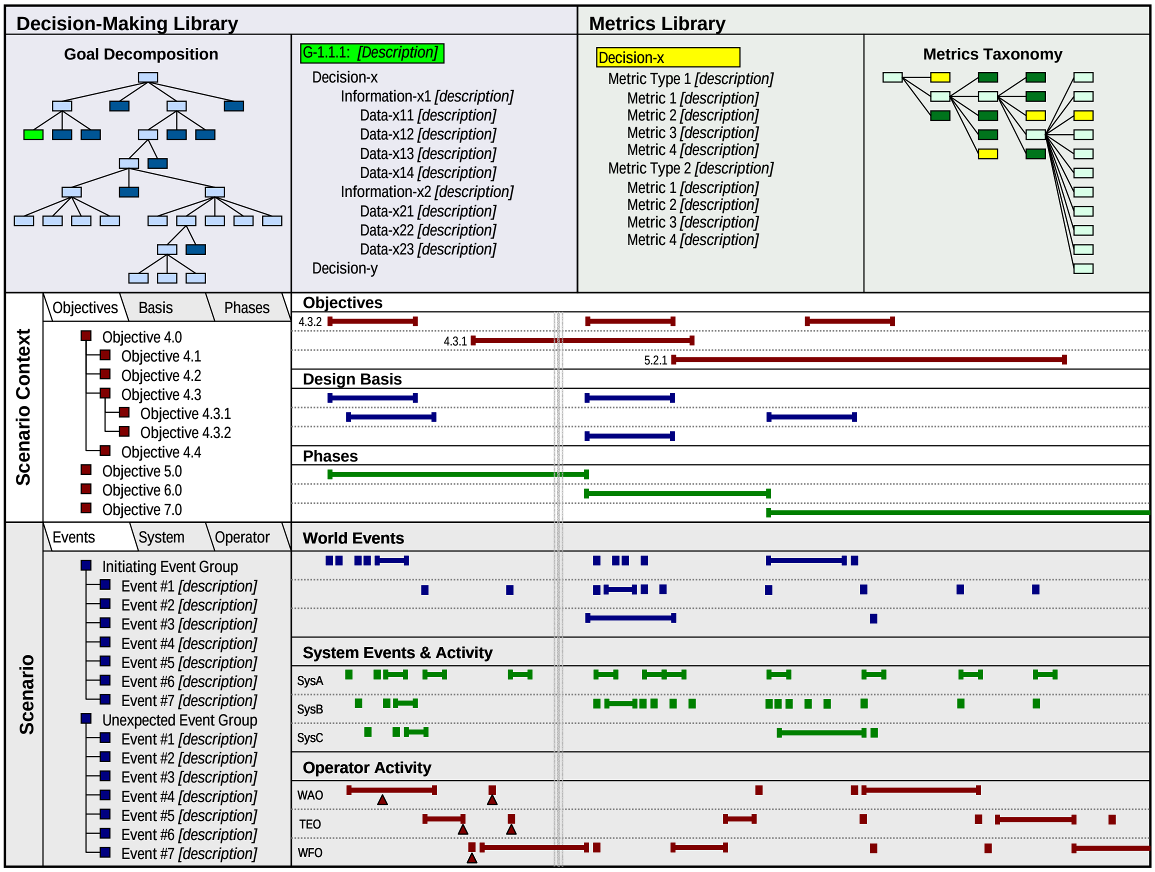


Figure 2. Scenario Management Tool Design Concept

a) Multi-layered Scenario Management:

At the heart of this tool is a Scenario Management Component that represents the scenario in three inter-connected layers (bottom third of the tool, depicted in Figure 3). These layers of the scenario depict inter-related events and activity that together define the critical characteristics of the scenario.

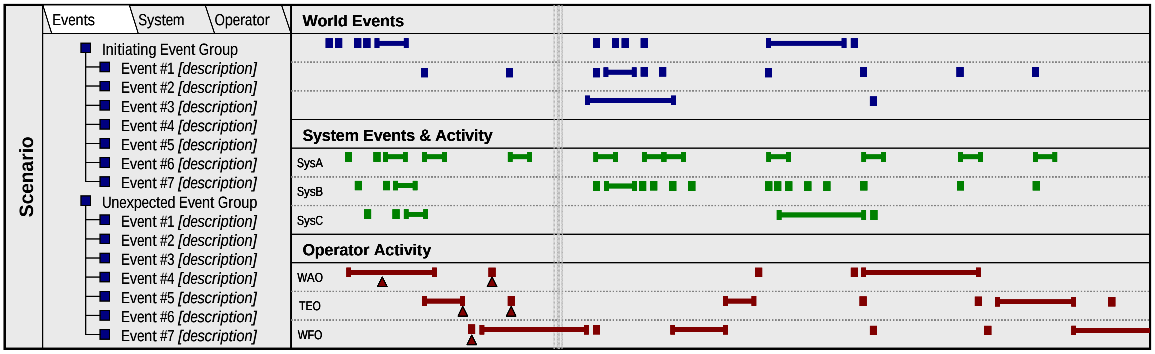


Figure 3. Scenario Management Tool Timeline Component

Below is a description of each layer:

- **World Events.**
    - Depicts events as either an instantaneous event or event duration.
    - Contains multiple sub-layers to depict physically or functionally distinct events (e.g., events from different adversaries).
    - Events are linked to the scenario context to capture the linkage from scenario phases and specific events.
    - Events are linked to exercise objectives to establish the mapping from objectives to scenario events.
    - Events are linked to System Events / Activity to depict the relationship between scenario events and system responses to these events.
    - Events are linked to Operator Activity (predicted and actual) to capture the critical relationship(s) between scenario events and the responses that they are designed to elicit.
    - Any dependencies between events (e.g., timing, sequence, etc.) are to be depicted within this layer.
- **System Events / Activity** (in response to the world events)
    - Depicts events and activity as either an instantaneous event or event duration.
    - System activity can include information provided to an operator by an information management / processing system.
    - Contains multiple sub-layers to depict physically or functionally distinct events (e.g., events from different systems / weapon platforms).
    - Events and activity are linked to World Events to depict the relationship between scenario events and system responses to these events.
    - Events and activity are linked to Operator Activity (predicted and actual) to capture relationship(s) and coordination between system and operator activity.
- **Operator Activity** (in response to world events and system activity)
    - Depicts activity as either an instantaneous event or event duration.

- Represents time windows available for successful decision-making, predicted operator activity, as well as actual activity (updated as the scenario unfolds in real-time). This actual vs. available / predicted serves as a key performance metric.
- Contains multiple sub-layers to depict distinct operator (e.g., activity by multiple members of the command team).
- Activities are linked to World Events to capture the critical relationship(s) between scenario events and the responses that they are designed to elicit.
- Events and activity (predicted and actual) are linked to System Events / Activity to capture relationship(s) and coordination between system and operator activity.
- Operator activities are linked to the scenario context to capture the linkage from scenario phases and specific decision-making demands / activities.

b) Scenario Context / Objectives:

The Scenario Context (middle section of the tool, depicted in Figure 4) captures three aspects of the scenario: (1) the scenario design basis (i.e., the underlying characteristics of the scenario that are meant to be captured by the specific events of the scenario), (2) the high-level description of the scenario phases that are designed to create the appropriate context for the objectives of the exercise, and (3) the mapping from exercise objectives to the scenario.

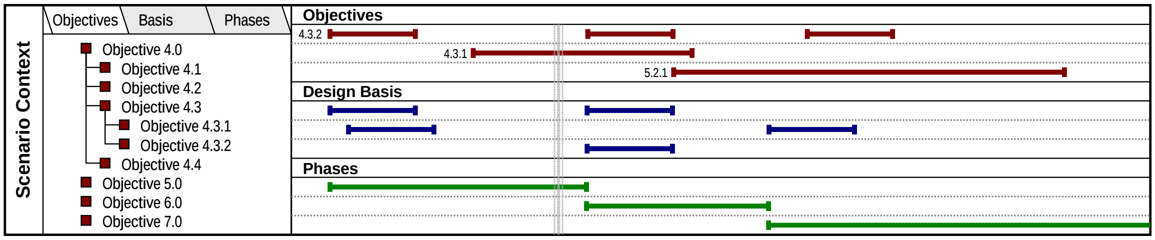


Figure 4. Scenario Management Tool Context Component

- **Scenario Design Basis.** This includes:
  - Individual and team **Knowledge and Reasoning Skills** that are an inherent part of the scenario. These can include known errors and biases (e.g., availability bias, concrete information bias). While humans are capable of processing complex information, known cognitive biases can hamper the decision-making process. For example, in information overload situations, decision-makers are prone to adopt mental strategies that simplify the problem, but which can lead to erroneous representations (Patterson, Roth, and Woods, 2001). The knowledge and reasoning skills are important to consider when developing scenarios because inclusion of event sequences meant to elicit these factors are an important characteristic of effective scenarios.
  - **Complicating Factors.** Scenarios should include not only routine or textbook cases, but also cases that challenge both individual cognition and

collaborative processes and reflect real complexities that arise in a domain. Complicating factors can be thought of as types of conditions that can arise in a domain to create opportunities for potential knowledge and reasoning skills to emerge.

These two sets of factors can then be combined to create a rich design basis for scenario development that then can be instantiated with mission-specific details that arise in the domain. Thus, the tool enables the specification of multiple specific scenarios for a given scenario context.

- **Scenario Phases.** In addition to applying cognitive processing and complicating factors to the scenario design problem, scenarios need to be credible and engaging to the participants. This means that the scenario needs to be carefully crafted to include relevant domain events and have sufficient realism. Typically, a number of scenario events are created that relate to each other and together form a particular phase of the scenario.
- **Exercise Objectives to Scenario Mapping.** This consists of two components:
  - **Objective Library / Selection.** Typically, objectives are maintained in a hierarchical structure, with high-level objectives decomposed into supporting objectives. Thus, selecting objectives to be addressed by the scenario can be accomplished in a similar fashion as with using tree graphs (e.g., file selection).
  - **Objective to Scenario Mapping.** One issue that will need some additional work is the determination of the best association between scenario events and objectives. That is, should the mapping be made to the scenario event or scenario phase or design basis? If the mapping is made at the higher levels, the linkage between scenario events and phase would result in all of the events within a phase being mapped to the objectives for that phase. If the mapping is made at the event level, the linkage between scenario events and phase would result in a phase being mapped to all of the objectives that the scenario events within that phase are mapped.

c) Decision-Making and Metrics Library:

In addition to defining the operator decision-making activity within a given scenario, an envisioned component to our tool is a “decision-making library”, which contains an organized repository of decisions collected throughout the use of the tool. This facilitates the use of decisions in subsequent scenario design efforts. Our design concept for this component is depicted in Figure 5.

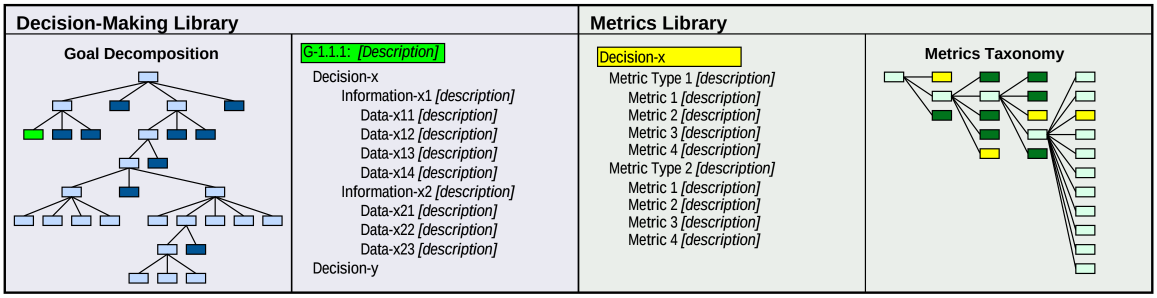


Figure 5. Scenario Management Tool Decision-Making Library Component

In addition to a definition of decision-making activities, we can identify and associate the information requirements for successful achievement of that particular decision-making activity. In this context, we define information requirements as the complete set of information elements necessary for successful resolution of the associated decision-making activity. Information requirements specify much more than specific data elements; it is data in context that becomes information (Woods, 1988; 1995). The data-to-information association is a potential extension to this repository; supporting data requirements can be associated with each of the information requirements to create the complete data-to-decision mapping.

Once we have defined the decision-making activity within a scenario as well as generically within the decision-making library, we can begin to define metrics for assessing the decision-making activity. Given the temporal perspective of the scenario management tool, the obvious first-level metrics involve the temporal aspect of the decision-making (e.g., time required for decision-making; decision time vs. LTIOV, etc.). It is important to note, though, that with the rich representation of the scenario (both with respect to the multi-layered representation of the scenario and the context), a wide variety of metrics can be defined (e.g., detection of the “garden path”; recognition of misleading information; identification of side-effects to decision-making activities). At a higher level, this framework supports the investigation of collaborative and macrocognition decision-making (due to the representation of multiple operators in the “Operator Activity” layer of the scenario) given the mapping from scenario context to events.

2. Conclusions

The critical issues identified through our investigation and design concept development include:

1. Practitioners use a variety of techniques to simplify the process of planning exercise scenarios.

Designing effective scenarios (to truly challenge the decision-making skills of the exercise participants) is a difficult task. As a result, most scenario designers make the process more manageable by:

- Building up a repository of scenario vignettes that can be adapted and modified to fit the specific situation.
- Simplifying the scenario events by only specifying the start time of the event.

- Not specifying the expected participant behavior.
- Having a loose connection between the scenario events and the exercise objectives.

2. The research community is constructing a basis for designing cognitively challenging scenarios

Critical to the effective exercise of cognitive skills is the design and conduct of challenging scenarios. As identified in this effort, this includes the specification of Knowledge and Reasoning Skills and Complicating Factors. These two factors constitute a robust basis for the design of cognitively challenging scenarios. We integrated these concepts into the Design Basis portion of the Scenario Context component of our tool.

3. Effective exercise planning and management requires the explicit representation of the linkage from scenario events to exercise objectives.

Our demonstration provides the ability to see the linkage from Exercise Objectives to Scenario Events or from events back to objectives. This is critical for the determination of achieving training objectives by the scenario. Establishing these links supports the ability to determine the significance of world events in exercising the training objectives. In addition, this linkage permits the planning and management of the training exercise to meet the desired objectives.

4. Scenario design requires the representation of the relationships between World Events, System Activity, and Operator Behavior

Fully described scenarios require the specification and coordination of system activity in response to world events. This includes automated system activity responding to the world events as well as information processing activity representing these events to the participants. Then, scenarios need to represent the expected and actual decision-making behaviors in response to the events and activities. Specifying expected behavior provides the ability to identify deviations from planned behavior as well as confirmation that the progression of operator activity is going as planned. Our design concept and demonstration provide a rich temporal representation of these relationships as a powerful mechanism for capturing the full specification of the scenario. This provides the exercise managers with the ability to (which is not apparent in the MSEL representation of the scenario):

- Detect deviations from planned scenario events (i.e., exercise entity not performing as expected; entity dropping off the network, etc.).
- Detect deviations from planned activity (i.e., operators failing to complete required actions within allotted time).

5. Scenario design requires the ability to specify different types of events to accurately depict the richness of the exercise.

Our design concept demonstration provides the ability to define and represent several different types of events:

- *Instantaneous Events.* If the event is truly instantaneous, it can be defined as such. However, the critical issue is that this is not the only way to define events (as in typical MSELs).
- *Duration Events.* Event start and end times can be defined to specify the duration of an event. This provides a much more realistic context for the definition of events.
- *Event Dependencies.* Event times can be specified based on a predecessor event, providing dependencies between events. This provides an extremely robust event context; especially for management of scenarios. With dependencies, if the predecessor event is delayed in the exercise, all the dependent events are also delayed.

6. Effective exercise planning and management requires a rich library of operator decisions and metrics for assessing decision-making effectiveness.

Our “decision-making library” component is a powerful support tool for increasing the effectiveness and efficiency at designing decision-centered exercises and evaluations. As the use of this tool grows, the repository of decisions and metrics will also grow, providing the ability to leverage the results from these prior efforts and construct a robust set of decisions and associated metrics for a given scenario.

Critical to the effectiveness of this decision library is an effective organizational framework for the storage and retrieval of the decisions. We propose to organize the decisions based on functional relationships based on the extensive work in Cognitive Work Analysis (CWA; Potter, Gualtieri, & Elm, 2003a; Bisantz et al., 2003; Roth et al., 2001; Vicente, 1999; Rasmussen, Pejtersen, & Goodstein, 1994). Therefore, the key aspect of our technical approach is the organization of decisions based on a functional model of the work domain. This framework also supports the investigation of collaboration and macrocognition given the grounding of the decision-making in naturalistic work domains.

To date, our Scenario Management tool has not been applied to any operational environments. However, millions of dollars have been spent over the past several years on developing advanced training systems, training infrastructures, simulations, and Live, Virtual, and Constructive (LVC) training system interoperability. While those systems and infrastructures bring a great deal benefit to trainees across domains, the success of a training exercise still depends on the ability of training managers to easily and rapidly digest the diverse, complex, and interacting streams of information that is inherent to large-scale, LVC distributed simulation-based training exercises and their ability to craft and conduct well-orchestrated scenarios within the training system. We believe that our Scenario Management tool is directly relevant and beneficial to training managers faced with those difficulties. To that end, we are targeting training managers for large-scope LVC environments, such as NAVAIR’s LVC training environment.

REFERENCES

Billings, C. E. and Woods, D. D. (1994). Concerns about adaptive automation in aviation systems. In M. Mouloua and R. Parasuraman (eds.), *Automation and Human Performance: Recent Research and Trends* (p.264-269). Hillsdale NJ: Erlbaum.

Bisantz, A. M., Roth, E. M., Brickman, B., Gosbee, L., Hettinger, L., & McKinney, J. (2003). Integrating Cognitive Analyses in a Large Scale System Design Process. *International Journal of Human Computer Systems,* 58177-206.

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.

Christoffersen, K. and Woods, D.D. (2002). How to Make Automated Systems Team Players. In *Advances in Human Performance and Cognitive Engineering Research,* Vol. 2, (pp. 1–12), Elsevier Sciences Ltd.

Dekker, S.W. & Woods, D.D. (1999). To Intervene or Not to Intervene: The Dilemma of Management by Exception. *Cognition, Technology and Work, 1,* 86-96.

Elm, W.C., Potter, S.S., Gualtieri, J.W., Roth, E.M., & Easter, J.R. (2003). Applied Cognitive Work Analysis: A pragmatic Methodology for Designing Revolutionary Cognitive Affordances. In E. Hollnagel (Ed) *Handbook for Cognitive Task Design.* London: Lawrence Erlbaum Associates, Inc.

Elm, W., Potter, S., Tittle, J., Woods, D., Grossman, J., & Patterson, E. (2005). Finding Decision Support Requirements for Effective Intelligence Analysis Tools. In *Proceedings of International Conference on Intelligence Analysis*. McLean, VA.

Feltovich, P. J., Hoffman, R. R., Woods, D. D., and Roesler, A. (2004). Keeping it simple: How the reductive tendency affects cognitive engineering. IEEE Intelligent Systems. IEEE Computer Society.

Hollnagel, E. and Woods, D. D. (2005). *Joint cognitive systems: Foundations of cognitive systems engineering.* Boca Raton, FL: Taylor and Francis.

Johnston, J. H., Cannon-Bowers, J. A., Smith-Jentsch, K. A. (1995). Event-based Performance measurement system. In *Proceedings of the Symposium on Command and Control Research and Technology,* June 1995 (pp 268-276).

Klein, G. (2007). Personal communication.

Lind, M. (1991). Representations and abstractions for interface design using multilevel flow modeling. In G. R. S. Weir and J. L. Alty (Eds.), *Human-Computer Interaction and Complex Systems*, London: Academic Press.

Lind, M. (1993). Multilevel flow modeling. *AAAI93 Workshop on "Reasoning about Function",* July 11-15, Washington, D. C.

Mumaw, R. J. & Roth, E. M. (1992). How to be more devious with a training simulator: Redefining scenarios to emphasize cognitively difficult situations. *1992 Simulation MultiConference: Nuclear Power Plant Simulation and Simulators*, Orlando, FL, April 6-9, 1992.

Patterson, E. S., Woods, D. D., and Roth, E. M. (1999). *Aiding the intelligence analyst in situations of data overload: A simulation study of computer-supported inferential analysis.* Institute for Ergonomics / Cognitive Systems Engineering Laboratory Report ERGO-CSEL-99-TR-02. Columbus, OH: The Ohio State University.

Patterson, E. S., Roth, E. M., and Woods, D. D. (2001). Predicting Vulnerabilities in Computer-Supported Inferential Analysis under Data Overload. *Cognition, Technology and Work,* 3(4), 224-237, 2001.

Potter, S. S., Roth, E. M., Woods, D. D., & Elm, W. C. (2000). Bootstrapping Multiple Converging Cognitive Task Analysis Techniques for System Design. In Chipman, Shalin, & Schraagen (Eds.), *Cognitive Task Analysis* (pp. 317-340). New Jersey: Lawrence Erlbaum.

Potter, S. S., Gualtieri, J. W., and Elm, W. C. (2003). Case studies: Applied cognitive work analysis in the design of innovative decision support. In E. Hollnagel (Ed.). *Cognitive Task Design.* New York, NY: Lawrence Erlbaum. Rasmussen and Lind, 1981.

Rasmussen, J. (1986). *Information Processing and Human Machine Interaction: An Approach to Cognitive Engineering*. New York: North Holland.

Rasmussen, J., Pejtersen, A. M., & Goodstein, L. P. (1994). *Cognitive Systems Engineering*. New York: Wiley and Sons.

Rasmussen, J. & Lind M. (1981). Coping with complexity. In H. G. Stassen (Ed.), *First European annual conference on human decision making and manual control*. New York: Plenum.

Roth, E. M., Mumaw, R. J., & Lewis, P. M. (1994). *An Empirical Investigation of Operator Performance in Cognitively Demanding Simulated Emergencies.* Washington D. C.: U. S. Nuclear Regulatory Commission. (NUREG/CR-6208)

Roth, E. M., Gualtieri, J. W., Elm, W. C., & Potter, S. S. (2002). Scenario Development for Decision Support System Evaluation. In *Proceedings of Human Factors and Ergonomics Society 46th Annual Meeting,* (pp. 357-361). Santa Monica, CA: Human Factors and Ergonomics Society.

Smith, P., Woods, D., McCoy, E., Billings, C., Sarter, N., Denning, R., and Dekker, S. Using forecasts of future incidents to evaluate future ATM system designs. *Air Traffic Control Quarterly,* 1998.

Smith-Jentsch, K. A., Johnston, J. H. and Payne, S. (1998). Measuring team-related expertise in complex environments. In Cannon-Bowers, J. A. and Salas, E. (Eds.)

Making decisions under stress: Implications for individual and team training. Washington, D. C.: American Psychological Association.

Vicente, K. J. (1999). *Cognitive Work Analysis: Towards Safe, Productive, and Healthy Computer-Based Work.* Mahwah, NJ: Lawrence Erlbaum Associates.

Vicente, K. J. & Rasmussen, J. (1992). Ecological Interface Design: Theoretical Foundations. *IEEE Transactions on Systems, Man, and Cybernetics,* SMC-22589-606.

Vicente, K. J. & Rasmussen, J. (1990). The Ecology of Human-Machine Systems: II. Mediating Direct Perception in Complex Work Domains. *Ecological Psychology,* 2207-249.

Woods, D. D. (1991). *Representation Aiding: A Ten Year Retrospective.* Charlottesville, VA: IEEE.

Woods, D. D. (1995). Towards a Theoretical Base for Representation Design in the Computer Medium: Ecological Perception and Aiding Human Cognition. In J. Flach, P. Hancock, J. Caird, & K. Vicente (Eds.), *An Ecological Approach to Human Machine Systems I: A Global Perspective.* Lawrence Erlbaum.

Woods, D. D. and Hollnagel, E. (1987). Mapping cognitive demands in complex problem-solving worlds. *International Journal of Man-Machine Studies, 26,* pp. 257-275.

Woods, D.D., & Patterson, E.S. (2001). How Unexpected Events Produce An Escalation Of Cognitive And Coordinative Demands. P.A. Hancock and P.A. Desmond (Eds.). *Stress Workload and Fatigue.* Lawrence Erlbaum, Hillsdale NJ.

Woods, D., Johannesen, L., Cook, R., and Sarter, N. (1994). *Behind Human Error: Cognitive Systems, Computers and Hindsight.* Crew Systems Ergonomic Information and Analysis Center, WPAFB, Dayton OH.

Woods, D. and Sarter, N. (1993). Evaluating the Impact of New Technology on Human-Machine Cooperation. In J. Wise, V. D. Hopkin, and P. Stager, editors, *Verification and Validation of Complex Systems: Human Factors Issues,* Springer-Verlag, Berlin.

Woods, D. D. and Hollnagel, E. (2006). *Joint cognitive systems: Patterns in cognitive systems engineering.* Boca Raton, FL: Taylor and Francis.