Enabling Engineering of Complex Systems Through Simulation-Based Experimentation

Zach Furness
The MITRE Corporation
McLean, VA
zfurness@mitre.org

Dr. Valerie Gawron
The MITRE Corporation
McLean, VA
vgawron@mitre.org

***Abstract* – We discuss lessons learned in the development and application of the Collaborative Experimentation Environment (CEE), a simulation-based environment for conducting multi-agency human-in-the-loop (HITL) experiments. Experiences across three different net-centric experiments (NETEXs) are summarized and the potential value of experimentation for improving multi-agency collaboration is reviewed.**

I. INTRODUCTION

Over the past decade, simulation-based experimentation has gained more and more acceptance as a viable means for uncovering problems in engineering of complex systems. Through immersing human "role players" into environments that can recreate many of the same stimuli that users would see when interacting with operational systems, interactions between systems and users in the context of real-world operations can be better understood. Such experimentation, done as early as possible in the design and development process, can identify problems in technology, policy, procedures, and training that can then be addressed prior to full-scale develop of the system.

However, the complexity associated with simulation-based environments make them difficult to use. The complexity comes from managing the integration of large numbers of simulations, systems, and support tools (databases, visualization capabilities, etc.). A large distributed simulation network that spans multiple sites and involves hundreds of users can take as long as a year to prepare for an event. Often times the insights that are gained in using these environments does not justify the investment in time and funds necessary to employ them.

At the other end of the spectrum are "table-top" exercises or seminar games in which participants are exposed to a variety of situations and scenarios and discuss potential solutions, but do not interact directly with operational systems in real time. While useful to help identify high-level gaps in plans and improve coordination among participants, these table-top exercises lack the fidelity needed to identify problems in interoperability among systems and potential gaps in procedures across different agencies.

MITRE's Collaborative Experimentation Environment (CEE) was developed to provide a "mid-level" simulation-based environment for executing experiments that provides enough fidelity to accurately represent the systems, processes, and critical elements of interest to government sponsors while minimizing the time necessary to employ such environments. The CEE is intended to focus on multi-agency experimentation to help identify gaps in technology, policies, procedures, and training across two or more government agencies. As the failures of 9/11 and Katrina pointed out, there is a growing need within the government for multiple agencies to coordinate their operations in response to various domestic threats, but the technologies, policies, procedures, and training that can improve coordination are still lacking.

One of the major tenets of the CEE is to build on the already established labs and infrastructure that currently exist within MITRE's centers supporting the Department of Defense (DoD), Federal Aviation Administration (FAA), and Department of Homeland Security (DHS). By linking together these existing labs and resources, we have been able to integrate existing systems and simulations that have already been used extensively within each center's programs to support their mission and leverage them for use across multiple sponsors with interconnected missions. We've been able to do this relatively quickly due to the adoption of a standardized architecture and a formal but flexible systems engineering process that is derived from experiences across all three centers.

The vision of the CEE is to quantify multi-agency mission effectiveness. While the creation of the environment is fundamental to achieving this vision, it is not the only element required. Partnerships with industry, academia, and other Federally Funded Research and Development Centers (FFRDCs), in addition to the government, are just as important in executing the experiments as provisioning of the environment. Over the past two years, MITRE has worked with numerous organizations including Aeropuertos y Servicios Auxiliares (ASA), Air Transport Association, International Association of Chiefs of Police, International Civil Aviation Organization, John Hopkins University Applied Physics Lab (JHU APL), Metropolitan Washington Airports Authority (MWAA), National Governors Association, NAVCANADA, Thales Raytheon Systems, Transport Canada, United Airlines, and US Airways, in use of the CEE. These associations have been crucial to ensure that subject matter experts and systems of record are part of the NETEXs that are conducted as part of the CEE.

II. NETEXs 08-01 and 08-02

Since 2007, numerous NETEXs have been conducted using the CEE. The first two, done in the spring of 2008, addressed issues related to air space security across multiple agencies including FAA, DOD, DHS, and the Department of Justice (DOJ). Specifically, these NETEXs focused on aspects of information sharing across the Domestic Events Network (DEN) – the 24/7 teleconference that was established during 9/11 and can involve dozens or hundreds of federal agencies at any one time. While the DEN has served as a useful means of facilitating communications among these agencies during a crisis, it is currently limited to voice-only communications. The two NETEXs 08-01 and 08-02 introduced several new technologies including a shared situational awareness tool as well as a chat application to augment existing voice-only communications.

The insights gained from these NETEXs were immediately used by the FAA, DHS, and other Industry partners to help re-evaluate procedures with respect to information sharing on DEN. Most importantly, the development and execution of these two NETEXs took approximately 3 to 4 months each – enabling these organizations to gain rapid insights into the problem without having to expend considerable time and effort to develop the experimental environment. Almost all of the major simulation and operational systems already existed either within MITRE, or through various industry partnerships. By relying on a standards-based architecture to the greatest degree possible, the necessary simulations, prototype applications, and operational systems were integrated quickly.

III. NETEX 09-01

NETEX 09-02, conducted in March of 2009, explored procedures, policies, and technologies of potential use in the prioritization of Unmanned Aerial System (UAS) missions across multiple agencies during a response to a Hurricane. Currently, UAS missions are limited to one flight per FAA facility over US Airspace, requiring agencies with competing mission requirements to rely on the FAA to make decisions with respect to mission priorities. This experiment evaluated an adjudication procedure in which Federal Emergency Management Agency (FEMA), would prioritize UAS missions across DoD, DHS, and the National Oceanographic and Atmospheric Administration (NOAA), prior to submission to the FAA. In addition, a situational awareness tool was introduced to the participants that gave them better visibility into mission nominations, assigned priorities, and available UAS assets.

Representatives from Air Force Northern Command (AFNORTH), Customs and Border Patrol (CBP), the FEMA, the FAA all participated in the experiment. One of the key findings of this experiment was that a "structured process and a central adjudicator would result in improved handling of UAS missions in the disaster response" [Maroney, 2009]. The NETEX helped to illustrate the "art of the possible" to the various organizations through a hands-on experience that enabled all of them to share similar perspectives on this issue. By doing so, it facilitated gaining consensus on a viable approach. The adjudicator procedure was implemented in time for the start of the 2009 Hurricane season –j ust 67 days after the completion of the experiment.

IV. NETEX 09-02

In August of 2009, the CEE was used to support a Pandemic Influenza Experiment (PIE) focused on improving the understanding of health screening of international passengers entering the North American continent during a Pandemic Influenza outbreak. The motivation for this experiment stemmed from proposed airport screening policies in which 100% of international travelers would be screened during a Pandemic outbreak. Coordination among the FAA, CBP, the local airports management authority, and the Center for Disease Control (CDC) is of particular importance as all these organizations have overlapping responsibilities. Additionally, international partners such as Mexico and Canada also potentially play are role in the management of re-routed flights and screening of passengers between North America and other countries. The focus of the NETEX was determining how information sharing capabilities could help those agencies improve coordination during such an event. That NETEX has already led to improvements in procedures between the FAA and CDC as well as follow up coordination meetings with ASA, FAA, and Transport Canada as well as the Air Mobility Command (AMC) and CDC.

V. DATA ANALYSIS

NETEX 09-02 offers several good examples of the type of data that can be collected in the CEE and the types of analysis that can be performed. The experiment focused on the use of a collaboration prototype that tracked mission requests, assigned priorities, available resources, and status of UAS missions in a way that was accessible to all participants. Three main vignettes were run during the experiment – the first with the collaboration tool and the establishment of a central adjudicator to prioritize mission requests, the second without the tool but still using an adjudicator, and the third without either an adjudicator or the tool. Without the use of the collaboration tool, participants resorted to email, phone, and chat which resulted in more messages being transmitted – a large portion of which related to the need to retransmit information. Figure 1 shows the average messages transmitted per mission need, while Figure 2 shows missing information and retransmission requests.

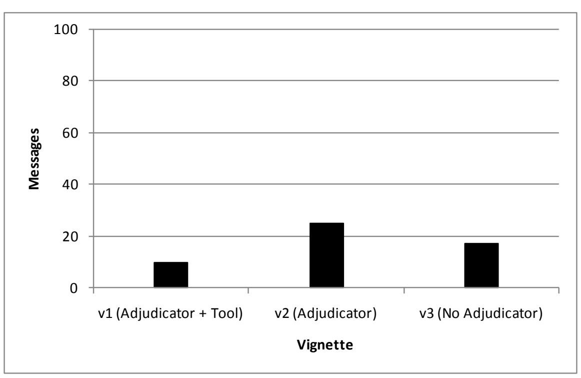

Figure 1. Average Messages Per Need

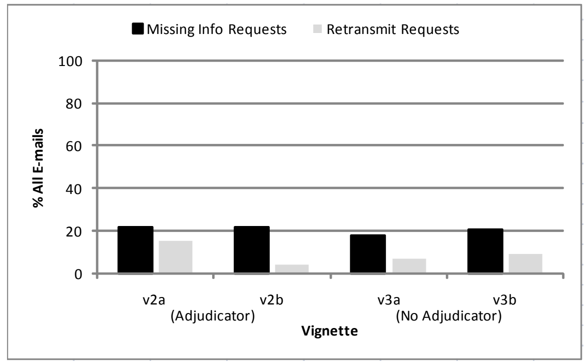

Figure 2. Number of Messages Requesting Missing or Previously Sent Information

In addition to the these types of system-level measures, NETEXs also rely on qualitative measures such as survey responses and observations of data collectors to complete the understanding of what happened during an experiment. Figure 3 below shows the results of participants to surveys administered immediately after the completion of a vignette that highlights their insights into the value of the information sharing tools they used. These results help to confirm that the lack of a collaboration tool hampered the efficiency of the participants during the experiment.

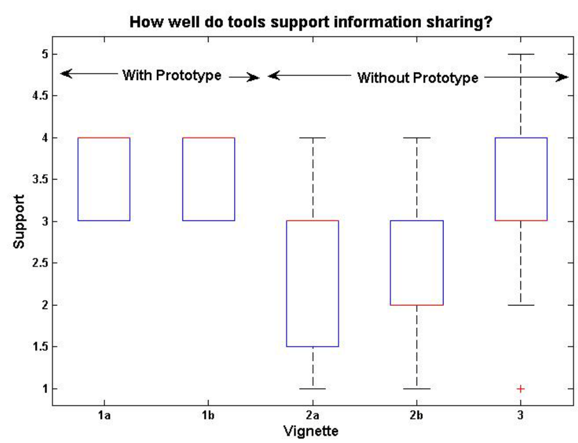

Figure 3. Information Sharing Survey Results

VI. OVERARCHING LESSONS LEARNED

Several excellent lessons have already come out of the CEE experience. Some of these are broad lessons that reflect the value of experimentation and some are specific to CEE's focus on multi-agency experimentation. Some of the major observations are described below.

A. Engagement Throughout the Process

While the experiment itself has substantial benefit to the participants, there is incredible value in the entire experimentation process. In the early stages of concept development and scenario definition, organizations gained a significant amount of insight into each other's procedures, plans, and responsibilities – well before the planned experiment. At the initial planning conference prior to NETEX 09-01, more than a dozen different government organizations attended including AFNORTH, CBP, FEMA, FAA, NASA, NOAA, Coast Guard, and National Guard to better understand the policies, procedures, and technologies that each already had in place. While this level of engagement is important to scope the experiment and ensure the proper focus, it has a secondary impact of bringing these organizations together to better understand their roles and responsibilities on multi-agency operations.

B. Simulation-aided Experimental Design

Early on in the experimental design of the PIE, it was difficult to come to consensus on the relevant measures for the experiment, and even harder to describe the importance of these measures to participating organizations such as the FAA, CDC, and CBP. Within a few weeks, the CEE team developed a process model that visualized the interactions among re-routed aircraft (to one of 19 designated quarantine airports), passenger arrival rates, and primary and secondary screening times. This model was key in providing a better understanding of the complex screening issues for both the participants and the experimentation team. Figure 4 below is an example of an artifact created to illustrate the flow of scenario events as part of NETEX 08-02.

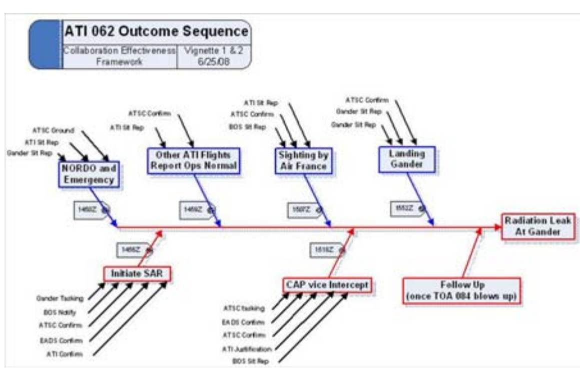

Figure 4. NETEX 08-02 Event Sequence Diagram.

C. Event Reconstruction and Forensics.

The ability to measure and quantify multi-agency effectiveness is one of the cornerstones of the CEE. But just recording responses in a simulation-driven experiment is not enough. Without the ability to take those measurements and use them to reconstruct the interactions that took place during the experiment, the value from the experiment will be limited. In a multi-agency experiment that mixes systems and HITL interaction, trying to do event reconstruction can be challenging. The CEE team devised an approach which not only captures data logs from webpages, operational systems, chat, and other collaboration tools but also correlates these data with recorded voice, and visual observations by data collectors. The data are then translated into a format that can be visualized in traditional timeline plots using such applications as SIMILE. The resulting mission threads can then be compared with graphical depictions of the scenario documented prior to the experiment to look for differences in the expected behaviors versus the observed. Figure 5 below is an example of such a timeline chart.

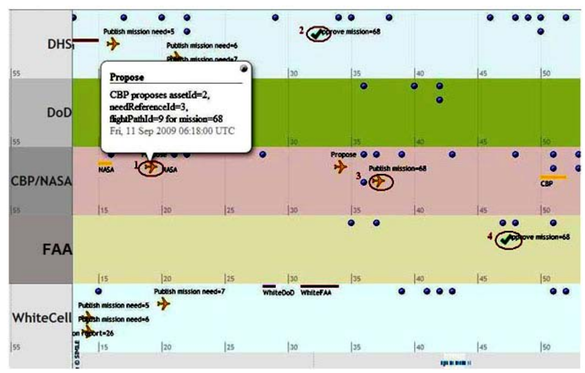

Figure 5. SIMILE Event Reconstruction Diagram

Other data collected includes web-based surveys that are filled out immediately after each vignette to ensure that participants' insights are immediately captured. The questionnaires help to document the reasons behind participants' decision making – often times not always easily inferable from their overt behavior. A web-based system provides a method for low-cost, rapid implementation of questionnaires that can be short, tailored to the participant's background, and provide easily retrievable and analyzable response data.

D. Disciplined Systems Engineering.

Being able to jump from and air security experiment to a UAS experiment and then to a Pandemic screening experiment would not have been possible with a well understood and documented systems engineering process. Implementing such a process is the key to being rapidly reconfigurable. Major elements of the process include the development of key artifacts through experimental design and planning, use of a standards-based infrastructure to help with integration of new simulations and services, and limiting design complexity to ensure that experiments can be completed in 3 to 4 months. Typical artifacts that are part of the systems engineering process includes technical and operational architectures, event-sequence diagrams, a master scenario event list (MSEL), test plans, and a data collection plan. The development of these, and other artifacts, is facilitated through internal Wiki and SharePoint services that are used to communicate updates.

E. Actionable Recommendations.

All the experimentation in the world is worthless unless it can provide actionable insights that participants can take-away with them and integrate into their daily operations. Once again, the experiment itself is not the only vehicle for establishing such insight. Vesting the participants in the design, planning, documentation, execution, and analysis of the experiment is crucial to help them understand key assumptions, and caveats that go into the experiment.

F. Measuring Situational Awareness.

Trying to quantify situational awareness (SA) of each of the participants during an experiment is extremely challenging. While the techniques outlined above are by no means a complete answer to the issue, there are some meaningful lessons coming out of the experience. First, instrument all the systems involved (email, chat, voice, web, etc.) including logging of data and time stamping of transactions. However, while this is necessary it is not sufficient to get a complete picture of SA. Coupling these data to observations from data collectors, surveys of participants, comparisons MSEL injects, are all equally important. Finally, being able to synthesize all of these data into a single coherent picture of what happened and why is paramount to gaining a complete understanding of the awareness of each participant and how that awareness contributes to the overall mission outcome.

VII. FUTURE PLANS

During 2010, MITRE plans to continue to execute NETEXs related to multi-agency air surveillance and UAS operations, in addition to expanding the capabilities of the CEE to address Mission Assurance and Cyber-Security experimentation. We are also expanding the functionality of the CEE to enable simultaneous classified/unclassified operations through the implementation of a multi-level security gateway between two different experimentation labs. Finally, we hope to establish an SDREN/DREN connection between the CEE and some of the major DoD experimentation facilities in the coming year to broaden the range of experimentation collaboration with government and industry partners.

VIII. CONCLUSIONS

The CEE is a rapidly reconfigurable distributed environment with three major uses: 1) understanding the effects of technologies, procedures, policies, and trainingon multi-agency operations; 2) generating quantitative data for empirically-based decision making on acquisition and policy; and 3) providing an environment for government and industry to explore concepts and evaluate technologies by leveraging existing internal and external resources to support multi-agency collaboration experiments. These resources include operational personnel, operational data, scenario generation tools, models and simulations, operational and/or emulated systems, and distributed interoperability infrastructure. The CEE vision is to quantify multi-agency mission effectiveness.

This paper discussed some of the experiences and lessons learned from simulation-based experimentation and the potential for using “mid-fidelity” simulation-based experimental environments for future engineering of complex systems. We feel that the direct involvement of users in such environments as early as possible in the acquisition cycle can have a profound impact on the successful development of these systems, especially in areas where these systems of systems are to be used across large, multi-agency enterprises.

REFERENCES

Maroney, D., et. al, “*HUREX – Experiment of Use and Prioritization of UAS in Hurricane Disaster Response*”, 9th AIAA Aviation Technology, Integration, and Operations Conference (ATIO) 21 - 23 September 2009, Hilton Head, South Carolina

Bigbee, A., Curtiss, J., Litwin, L., Harkin, M., “*Models in Multiagency C2 Experiment Lifecycles: The Collaborative Experimentation Environment as a Case Study*, to appear in The International C2 Journal, 2010.