Cyber Red/Blue and Gamified Military Cyberspace Operations

Nancy L. Crabtree and Joshua A. Orr

Lincoln Laboratory researchers designed a serious game to investigate how such games could aid cyber security specialists in developing and practicing cyber defense strategies. Proof-of-concept experiments conducted with the prototype Cyber Red/Blue game yielded insights into game design and player behavior. An improved understanding of game dynamics can inform games' future development as tools for cyber security research, training, and real-world mission applications.

The security of the cyber domain has grown rapidly into a major concern for the U.S. government and American society in general. The Department of Defense, National Security Agency, and Department of Homeland Security are working actively to ensure that the proper protections, situational awareness, decision support, and information-sharing mechanisms are in place to protect the U.S. critical infrastructure, including data, against major cyber attacks.

To support these government agencies in improving the nation's ability to withstand cyber attacks, MIT Lincoln Laboratory's Cyber Security and Information Sciences Division developed the Cyber Red/Blue serious gaming platform and defense-oriented game to explore the potential benefits serious gaming may provide for cyber security and to learn more about the human role in cyber defense. Cyber Red/Blue leverages the Laboratory's red team (offense) versus blue team (defense) exercise approach to explore the effectiveness of techniques and systems designed to respond to threats.

Key Aspects of the Cyber Domain

The cyber domain is an evolving human-made area of science, engineering, and practice that encompasses the hardware, software, networks, and data that drive the processing of information and the functioning of software-assisted physical devices. Because the cyber domain is human-made, many of its security challenges are different from those of the physical sciences. The rules of cyber operation can change rapidly, unlike the laws of the physical domain. Complexity in the cyber environment grows continuously, spurred by the adoption of new technologies and the ever-changing characteristics of the data these new technologies produce.

Human interaction with computers—the "human in the loop"—plays a critical role in the realization of cyber security goals, but this role is not well understood. Researchers working in cyber security need to gain a better understanding of not only where cyber security risks lie but also how humans can engage to minimize those risks.

Five key considerations for exploring human behavior in the dynamics of cyber security and operational resilience are the enterprise mission, the cyber threats to that mission, the mission-enabling infrastructure against which attacks occur, the human defenders' operational processes, and the roles that humans play in cyberspace operations. Central to these considerations is an understanding of the attack surface, which is understood as all the points at which a cyber attacker can gain access to a computer system or network.

Addressing Key Challenges

Two major areas in which serious games and gamification (the application of game-like elements to non-game activities) could enhance cyberspace operations are in the reduction of information ambiguity, often referred to as the fog of war, and the decrease in the time T to observe, orient, decide, and act (T_{OODA}) with respect to one's adversary.

Fog of war is a term used by the military to describe an operational situation in which unclear information leads to ineffective and/or inefficient decision making. Carl von Clausewitz in his 1832 book *On War* coined the term fog used in this manner and illustrated its attributes as follows [1]:

> ...[A] general in time of war is constantly bombarded by reports both true and false; by errors arising from fear or negligence or hastiness; by disobedience born of right or wrong interpretations, of ill will, of a proper or mistaken sense of duty, of laziness, or of exhaustion; and by accidents that nobody could have foreseen. In short, he is exposed to countless impressions, most of them disturbing, few of them encouraging....

John Boyd, a colonel in the U.S. Air Force, described the concept of the OODA loop in a number of briefings on military strategizing. In the most often quoted of these, delivered in 1986 [2], he said that "...in order to win, we should operate a faster tempo or rhythm than our adversaries—or, better yet, get inside [the] adversary's Observation-Orientation-Decision-Action time cycle or loop." Today, the OODA concept is widely used as a means to distill tasks into these four basic components in the study of decision making and the design of decision support systems, making it a concept central to the Cyber Red/Blue serious game.

Conflict in cyberspace, while new and technically challenging, still conforms to traditional models of conflict. As do defenders of other domains, defenders of cyberspace strive to minimize the fog of war and T_{OODA}, either deliberately or intuitively. However, the volume, velocity, and variety of operations in the cyber domain, coupled with enormous attack surfaces and the low cost to adversaries of mounting a cyber attack, make the goal of minimizing both information ambiguity and T_{OODA} very difficult with the tools available. The findings, training applications, and user interface improvements made through serious games and gamification research have the potential to greatly decrease fog of war and T_{OODA} while increasing operational efficacy in cyberspace.

Benefits of Serious Games

Cyber Red/Blue explores the idea that serious games can benefit practitioners, operational planners, and researchers of cyber security in the following ways:

- As game players, cyber security practitioners can master tools and processes through experimentation in a safe learning environment.
- Planners can think through scenarios to realize the dependencies, potential interactions, and available courses of action the game players face.
- Planners can observe gameplay and evaluate measured results of actions to gain insights that enable them to rapidly test and refine plans in a simulated environment before enacting those plans on the cyber "battlefield."
- For researchers, serious games can provide a methodology, a controlled environment, and iteration capabilities that allow them to isolate and measure aspects of cyberspace operations.

Employing game design elements into cyberspace operations' "battle management" systems may also improve human capacity to manage complex cyberspace operations. In the future, lessons learned from data collected in exercises using serious games could rapidly inform new mechanics for gamified operational counterpart systems, much like beta testing new game elements in precise market segments informs general-availability releases of personal computer games.

Cyber Red/Blue: The Platform

Cyber Red/Blue consists of a playable simulation platform and an initial prototype game. The platform offers an instrumented interface, a configurable simulated enterprise computing infrastructure, and a tool to create different game scenarios to allow human defenders to practice against automated cyber attackers in a measurable environment. Example configuration elements include network topology, the capabilities and numbers of workstations and servers in the enterprise, and courses of action that are available to players. Different game scenarios can include, for example, different kinds of cyber attacks, the incorporation of actual enterprise data, and tips and cues available to players.

The platform provides modular and extensible software models that execute predefined actions in response to player interactions with the simulated enterprise computing infrastructure environment and to player commands. The models interface with a publish/subscribe–based discrete-event simulation engine to enable a dynamic response to player actions by the simulated attacker and simulated enterprise infrastructure, and to generate recordings of the game events. Cyber Red/Blue includes emerging decision support tools that can be integrated within a unified cyber incident commander workflow.

Cyber Red/Blue: The Game

The prototype of the Cyber Red/Blue game was designed inside the platform as a defensively focused game in which the blue roles of planner and player defend against a simulated red attacker. The game addresses some of the cyber security decision support challenges of the enterprise defender in an operational environment.

During the initial experimental trial, players were presented with a fog-of-war problem: protect an enterprise environment while sifting through increasingly voluminous datasets. Players were required to interpret and respond to a large number of available logs and alerts generated by the enterprise's different computer systems in order to find the "needles in the haystack" that represented credible threats. Players applied an understanding of the situations presented to them to evaluate potential courses of action and to select the most appropriate action to initiate additional protections for the enterprise environment.

As players and planners made decisions in the game, the simulation responded, resulting in changes to the remainder of the gameplay. The combinations of player responses had impacts on the ability of the simulated operational infrastructure to support the enterprise mission. Impacts can include changes to the confidentiality, integrity, and availability of data and services, and the automated attacker's likelihood of taking control of the operational environment. For example, the players' ability to detect cyber attacks through their situational awareness capabilities directly correlated to their subsequent ability to respond to these attacks and take appropriate courses of action to prevent future attacks. These first-level impacts culminated in changes to the state of the enterprise mission.

The different aspects of gameplay were mapped to the different elements of the OODA loop process to give us a deeper understanding of the human needs in each of those areas. For example, situational awareness actions were mapped to the *observe* and *orient* elements of the OODA loop. Courses of action were mapped to the OODA loop *decide* and *act* elements. More details on these aspects are described in the later section on human-machine interface and displayed in Figure 2.

Developing Cyber Red/Blue

The development approach for Cyber Red/Blue was divided into three main phases: (1) survey existing simulation capabilities, (2) apply the survey findings to the design and construction of the platform, and (3) use the platform to create and run a game that has an instructive scenario.

Analysis of Pre-existing Capabilities

In our initial step, we surveyed six existing human interaction–based simulation approaches and graded each on four categories: focus, scope, responsiveness, and scaling cost. Note that in the survey *technical defense* refers to measuring the effectiveness of the computer defenses themselves (such as access controls, software

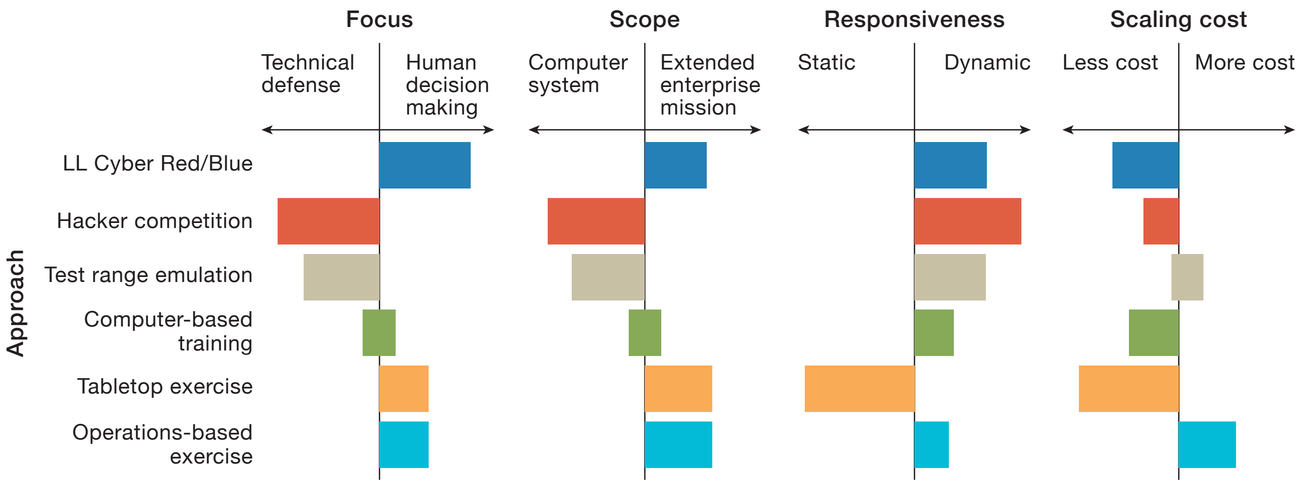


FIGURE 1. This comparison of the six simulation approaches listed on the left shows the advantage each has in the four categories listed across the top. Each category is divided into its contrasting characteristics, and the width of a colored bar indicates the relative level of advantage.

and hardware configurations, or software algorithms). In contrast, *human decision making* refers to measuring the effectiveness of the strategic and tactical approaches chosen by human decision makers (such as mission commanders responsible for making risk decisions and tasking resources at key points during the red/blue scenario). The summary findings are displayed in Figure 1.

Cyber competitions such as Capture the Flag train analysts, developers, and system administrators in a highly dynamic, emulated real-world environment through a deep emphasis on the elements of technical defense required at the computer system level. Monetary costs can be relatively low per simulation exercise instance. Computer test ranges, such as the Department of Defense National Cyber Range [3], consist of computer virtualization platforms that can be used to evaluate technical defenses of computer system interactions in a dynamic but controlled environment. Test ranges provide greater scaling capabilities than Capture the Flag but at the increased cost of a dedicated emulation environment.

At the time the survey was taken, a number of computer-based training resources were found that were oriented toward fulfilling certification and compliance requirements, and the list has expanded to include a number of online courses, such as SANS training [4] and the Department of Defense Cyber Awareness Challenge Training at Fort Gordon, Georgia [5]. Tabletop exercises emphasize the human decision-making processes of teams, but these exercises do not provide a quantitative measurement of those processes. Live operations-based exercises that make use of master scenario event lists can provide a high level of technical and decision-making realism, allowing for the wide scope of the extended enterprise mission and some dynamic outcomes, but these benefits come at significant system cost and complexity.

Developing a Needs-Based Capability

Cyber Red/Blue was designed to provide qualitative and quantitative measurement capabilities for human decision making in the context of a defensive cyberspace operation, but on a smaller scale and significantly leaner budget than the scale and budget of live operations–based exercises. The agility and low cost of the Cyber Red/Blue platform gives researchers and planners additional opportunities to experiment at more frequent intervals.

Cyber Red/Blue consists of four basic elements that are categorized as either human or automated computing components (Table 1):

1. Human-in-the-loop element. The human game players act as defenders working in a team to break the attacker's *kill chain* (i.e., a sequence of actions leading up to and including an attack). Through the game console's graphical user interface, players use simulated tools to make decisions and take actions.

Table 1. Cyber Red/Blue Simulation and Game Elements

HUMAN COMPUTING COMPONENTS		AUTOMATED COMPUTING COMPONENTS	
Human-in-the-loop element	**White cell element**	**Automated attacker element**	**Automated cyber activity element**
Attempts to break the stages of the kill chain	Develops objectives for game	Executes the stages of the kill chain 1. Undergo staging and reconnaissance 2. Gain access 3. Develop targets 4. Deploy attack 5. Verify, assess, persist in attack	Simulates enterprise environment
Decides, acts, observes, orients, as part of human-machine interface	Observes players and offers mentoring	Responds to player actions dynamically	Provides technical feedback
Utilizes technology tools to determine situational awareness, decision support, courses of action	Analyzes player activity		Creates smaller threats for game

This table summarizes key game role elements of the Cyber Red/Blue simulation. The human-in-the-loop element represents the actual game players defending the enterprise mission and its computer infrastructure. The automated attacker element is the software developed to run on the Cyber Red/Blue simulation platform that automatically executes attacks against the mission and infrastructure. The white cell element represents human analysts responsible for setting and assessing exercise outcomes. The fourth role element is automated cyber activity, which is software developed to run on the simulation platform to automatically execute the enterprise mission and its associated enterprise infrastructure background traffic.

2. Automated attacker element. The simulated attacker executes a prescribed kill chain to reach predefined objectives and is able to respond dynamically to player actions.
3. Automated cyber activity element. Configurable automated network traffic simulates the traffic of the enterprise environment that the human game players are working to protect. Through updated situational awareness indicators on the human-machine interface, this element also provides game players with feedback to inform future decisions. Additional background traffic simulates the multiple activities that can be observed in the enterprise cyber environment.
4. White cell element. Analysts responsible for setting and assessing exercise outcomes work with game planners to develop the exercise objectives. They then observe and analyze player activity to ensure the objectives are being met.

Human-Machine Interface

The human-in-the-loop element interacts with and plays the game via a role-based human-machine interface console. Figure 2 depicts the initial console layout. We did not undertake to develop a novel user interface, but rather we wanted to simply build an interface that would allow interaction with the simulated environment such that metrics could be collected. The key concept for the reader to take away from this figure is the mapping between OODA activities and potential player actions, and identification of additional tools that support evaluation during and after the game.

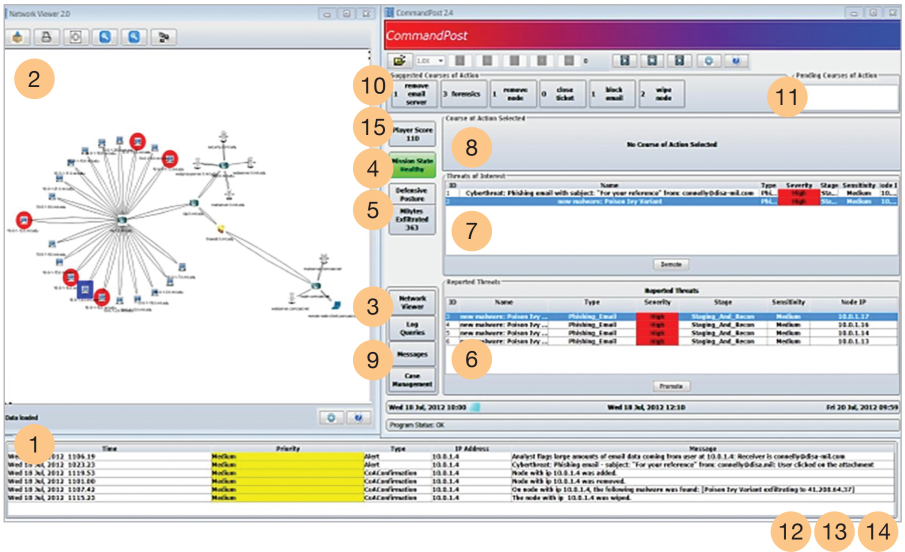
For the prototype game displayed in the figure, one example of game play function is (2) Network Display, a representation of an operational tool used for enterprise infrastructure situational awareness. The Network Display panel gives game players a diagram of the enterprise infrastructure configured for the game and updates dynamically to depict the fluctuating state of the enterprise computer infrastructure on the network. Lines depict network connections, and circles represent computers on the network. Red outlines indicate computers that have been attacked by the automated attacker, and blue outlines indicate computers that have had defensive courses of action taken on them by the game players.

Other operational capabilities displayed in the different panels include the following examples:

- Player tipping and cuing hints (e.g., Intelligence) for situational awareness provided by (1) Message Panel
- A list of identified threat types used to orient players to the mission threat environment in the context of key mission functions and guide them toward potential defense decisions, as provided by (6) Threat Context Panel
- A list of potential player courses of actions (CoAs), each with an explanation of their preconfigured risks to the mission and potential defensive contributions, as provided by (10) View CoA Costs and Benefits

Gameplay for Decision Support Challenges

One current decision challenge in the cyber domain is caused by the rapid escalation of threats. Daily, defensive operators and decision makers must parse copious amounts of uncorrelated data to find nontrivial pieces that can lead to the identification of ongoing threat activity. At the same time, information necessary to balance mission and security may be unavailable because of an incomplete understanding of the different cyber components on which the enterprise mission depends. This inefficient production and consumption of situational awareness information enables adversaries to rapidly evolve and intensify their activities without being detected when they are active, causing defenders to identify threats mostly post mortem. Once an incident is identified, decision makers must synthesize available information quickly to contain and remediate the threat and at the same time minimize mission impact. In other words, the adversary can observe, orient, decide, and act faster than today's defenders can. Attackers need only focus on their area of interest while defenders must be vigilant across the entire cyber mission



Threats of Interest

ID	Name	Type	Severity	Stage	Sensitivity	Node I
1	Cyberthreat: Phishing email with subject: "For your reference" from: connelly@disa-mil.com	Phi...	High	Sta...	Medium	10...
2	new malware: Poison Ivy Variant	Phi...	High	Sta...	Medium	10...

Reported Threats

ID	Name	Type	Severity	Stage	Sensitivity	Node IP
3	new malware: Poison Ivy ...	Phishing_Email	High	Staging_And_Recon	Medium	10.0.1.17
4	new malware: Poison Ivy ...	Phishing_Email	High	Staging_And_Recon	Medium	10.0.1.16
5	new malware: Poison Ivy ...	Phishing_Email	High	Staging_And_Recon	Medium	10.0.1.14
6	[illegible]alware: Poison Ivy ...	Phishing_Email	High	Staging_And_Recon	Medium	10.0.1.13

Time	Priority	Type	IP Address	Message
We[illegible]012 1106.19	Medium	Alert	10.0.1.4	Analyst flags large amounts of email data coming from user at 10.0.1.4: Receiver is connelly@disa-mil.com
Wed 1[illegible] 2012 1023.23	Medium	Alert	10.0.1.4	Cyberthreat: Phishing email - subject: "For your reference" from: connelly@disa.mil: User clicked on the attachment
Wed 18 Jul, 2012 1119.53	Medium	CoAConfirmation	10.0.1.4	Node with ip 10.0.1.4 was added.
Wed 18 Jul, 2012 1101.00	Medium	CoAConfirmation	10.0.1.4	Node with ip 10.0.1.4 was removed.
Wed 18 Jul, 2012 1107.42	Medium	CoAConfirmation	10.0.1.4	On node with ip 10.0.1.4, the following malware was found: [Poison Ivy Variant exfiltrating to 41.208.64.37]
Wed 18 Jul, 2012 1115.23	Medium	CoAConfirmation	10.0.1.4	The node with ip 10.0.1.4 was wiped.

Situational awareness
(Observe, orient)
1 Message panel
2 Network display
3 Log inquiries

Decision support
(Orient, decide)
4 Mission health panel
5 Threat level
6 Threat context panel
7 Threat of interest
8 Suggested course of action
9 Case management report

Courses of action (CoAs)
(Decide, act)
10 View CoA costs and benefits
11 Act by executing

White team support
12 Replay
13 Pause
14 Game/visualization control
15 Score

FIGURE 2. This figure depicts the prototype Cyber Red/Blue human-machine interface, which allows a game player to use multiple gameplay panels to execute steps from the OODA loop during the game—observe, orient, decide, and act. These OODA steps can be mapped to the different player functions, described at right of the graphic, to enable researchers to analyze player actions during and after gameplay. The mapping approach allows new gameplay panels to be swapped into different games without requiring the underlying analysis and measurement approach to change.

area to be defended and must understand how a cyber threat translates into a mission impact.

We prototyped and deployed our first game in the Cyber Red/Blue platform to test the ability of the simulation to measure human-in-the-loop capabilities while executing a game scenario. Specifically, we focused on the ability to measure game metrics related to operator environment and tools in order to understand how the environment and tools affect decision makers' fog of war and their ability to observe, orient, decide, and act. The goal of this game was to enable researchers to probe three basic questions:

1. How do various human-perceived observable artifacts (i.e., email logs, malware alerts, phishing tips, network topology, and system state) impact fog of war and T_{OODA}?
2. How do various technologies, in the form of simulated tools for situational awareness, decision support, and available courses of action, impact fog of war and T_{OODA}?
3. How do various stimuli, in the form of interactions, impact fog of war and T_{OODA}?

We tuned the Cyber Red/Blue platform to measure human-in-the-loop responses to observable artifacts by automatically tracking players' use of simulated defender tools (measured through player input to the user interface) and timing between automated stimuli and player response (measured by capturing timestamps for each event). To complement the quantitative measurements made within the system, the human analysts, i.e., the white cell element, were capable (through direct observation during the game and automatic replay of screen actions after the game) of identifying additional qualitative nuances in human perception capabilities, internal knowledge, player biases, and other psychological factors.

Initial Gameplay

For our initial game, we configured a simulated network topology consisting of server and workstation nodes on an enterprise local area network (LAN) connected to the Internet via a firewalled router. User nodes and servers executed the enterprise mission by passing email messages between themselves. The player console for the blue defender was simulated to reside on the enterprise LAN to monitor and protect the organization. The red attacker's simulated location was outside the enterprise within the Internet.

During the game, the automated attacker element simulated the red actor sending phishing emails with malicious content to blue enterprise clients. As the game progressed, some computers within the blue defender's area of responsibility were infected by the phishing emails, as represented by the red circled nodes in Figure 2. Once infected, a computer began sending out its own phishing emails and eventually started to exfiltrate mission data to the attacker.

The red attacker's goal in the game was to exfiltrate data from as many nodes as possible and to compromise the networked infrastructure by infiltrating enterprise servers from established footholds on blue enterprise nodes. Simulated attacker success meant the attacker would be capable of controlling the confidentiality, integrity, and availability of mission services. The human player's job as defender was to identify these attacks and mitigate their effects.

Throughout the game, players were presented with many observable artifacts, including a number of threats. Upon *observing* these artifacts, players could choose to "promote" threats (raise them to a higher monitoring priority level) when players determined the threats were of highest risk to the enterprise and mission. Phishing-email alerts were presented as a central threat, and players were notified of infection when nodes in the network viewer were highlighted red.

Players worked to *orient* themselves and determine the scope of the threat by performing a log query to identify other infected nodes. When players discovered 10 more infections, they decided to promote the threat. Once the threat was promoted, the player was able to view suggested courses of action and *decide* which of the actions was the most appropriate next step. To help guide gameplay, each course of action had a description of the costs and benefits to taking it.

One course of action option was to escalate the enterprise threat level, much as U.S. Armed Forces' Force Protection Conditions are elevated in response to potential threats to the nation. Other options were to block email containing specific characteristics so that the enterprise could be protected from future attacks of the same type or to remove a node from the network so that it could not communicate with other computers. Players could also simulate forensic investigations on computers to determine the underlying states of the computers and to gain an understanding of a particular threat. Additional options included wiping a node to remove malware and bringing nodes back online.

Course-of-action selection and implementation invokes some level of impact to the enterprise mission. For example, changing the enterprise threat level also changes the available courses of action. Blocking email that exhibits specific characteristics decreases the amount of email traffic exfiltrated and effectively decreases threat level, turning the mission's status panel to green to indicate mission integrity. Taking an infected email server offline stops all mission traffic and decreases mission health, indicated in red (serious mission breakdown) on the mission health panel.

In one game instance, players reviewed the infected node but did not block any additional email traffic. Because players did not choose that course of action, additional nodes became infected and the attacker exfiltrated mission data. The mission health panel changed to yellow to reflect a moderately compromised mission. After that, players had to scramble to keep up with the new level of threats. Eventually, mission health went to red because an email administrator became infected from the same phishing campaign and infected the email server.

Gameplay Findings

We played several games with separate teams of cyber researchers, security personnel, and decision commanders. We sought to create a baseline for future evaluations of decision support tools and human decision behavior, to gain feedback for improving the platform and presentation of decision support tools, and to provide insight on useful scenarios and exercise objectives.

To measure results, we first prepared the game environment by generating observable artifacts that could be measured as separate events, including operational email logs, malware alerts, and phishing tips. We configured network topology and made prototype tools available for players to monitor and control player actions during the different OODA steps. We configured the prebuilt attacks that the automated attacker would execute and the prebuilt courses of action that would be available to players at each enterprise threat level.

Before each game, we configured separate automated attack game scenarios. Each game scenario included the same enterprise and mission data, but we reconfigured the speed at which the attacks occurred to be slower at each consecutive game and increased the number of alerts that were generated in response to each attack. The consecutive game changes were necessary to allow the game players to work through the game scenario within a one-hour period.

Using instrumented results and white cell observations, we made two key findings. First, players spent most of their time on the orientation step, attempting to understand the elements of the log query tool to identify correlations between the threat context information and log query results. Second, player feedback focused on how the tools could be enhanced to improve results. Players' suggestions included adding proactive defensive capabilities to increase the security of enterprise operations before attacks occurred and enhancing the game tools to allow players to better understand attacks as they unfolded.

These findings led to several useful lessons learned:

1. "Train like you fight." We learned that for cyber serious games to be useful for practicing attacker scenarios and learning training objectives, it is important to provide the same tools and cyber environment players will face in the operational environment. In their game assessments, players focused on how the tools helped them play the game. Many recommendations from the game players related to improvements in the usability of different game console elements. This kind of feedback would be useful if we were seeking to evaluate real tools under development; however, because our tools were merely constructs intended for gameplay only, this attention to the tools diverted players' feedback from the game itself. When a game tool does not have the accuracy to emulate the real-world tool, it does not provide for the development of "muscle memory" for specific tasks, and presents the further risk that conceptual tools might inadvertently teach players the wrong lesson. These observations confirm the benefit of providing pluggable frames for inserting tools players would use in a real operational environment, especially if the game has a training objective.
2. Orientation. From our game results, it appears that without the right human decision support tools,

orientation can be the most time-consuming phase of the OODA loop in the cyber domain. Our instrumented game components allowed us to make comparisons between the times spent on the different phases of the OODA loop in order to come to that conclusion. We noted that players did not spend much time on the observe activities, such as network topology relationships and the in-depth node information available from the network viewer or from the out-of-band messages screen that collected miscellaneous enterprise information from different sources. Instead, during most of the game, players concentrated on the orient activities. Once they were oriented, players went quickly to the decide and act stages. Case management data confirmed this result. White cell members were able to observe conversations between team members to confirm that players spent the most time attempting to correlate the underlying sequence of events and did not devote much time to comparing potential course-of-action strategies for responding to the threat. Because players were viewing real operational logs but using a conceptual log correlation tool, it would be useful to perform further comparisons with real operational tools to identify the impact tools can have on condensing the orientation phase to speed up T_{OODA}.

3. Deep insight. We found that basing the game on the Lincoln Laboratory red versus blue concept could give us a multifaceted understanding of cyber decision-making processes. Our approach—which uses observable artifacts, the unified workflow, and simulated cyber models—measures multiple dimensions of player behavior simultaneously; it also provides a basis for comparing between operational tools and underlying assumptions to gain a better understanding of their impacts on defenders' success in managing challenges, such as decreasing fog of war and $T_{\text{OODA.}}$

The level of abstraction was sufficient to allow players to initially track and respond to threats. While the tools were not accurate representations of specific real-world tools, they were accurate enough to reveal the lack of correlation between different cyber technologies available at the time the tests were run and the effect of this lack on the time needed to orient.

As a result, players offered a number of useful suggestions to address this lack of correlation between information elements. These suggestions included adding summarized metadata to tie system names and IP addresses back to their users, making the dependencies between the mission functions and cyber systems involved explicit, and providing transparency as to how mission health levels, costs, and benefit calculations were made. The right kind of platform instrumentation to measure human behavior on real and candidate tools, and its use to execute a game scenario and submit player feedback, could lead to a serious game (or gamification using applied serious gaming concepts) that can provide a useful format for measuring training results and evaluating the effectiveness of cyber and human tools.

The first two lessons largely validate, in a game environment, concepts that continuously plague the operational community, while the third highlights an opportunity previously unavailable and uniquely plausible in the cyber domain. How, then, might serious games begin to address these issues?

Looking Forward: Gamified Military Cyberspace Operations

Serious games like Cyber Red/Blue provide both a controlled game-like venue to answer specific experimental questions and a training sandbox. Gamification can take concepts out of the sandbox and into the operational world in hopes of achieving higher efficiency and effectiveness through "the application of game design principles in non-gaming contexts" [6]. Let's look at how gamification, informed by serious game experimentation, can begin to address these findings toward decreasing fog of war and T_{OODA}.

Train Like You Fight

Training like you fight, a concept fostered in military doctrine, leads to a soldier's development of procedural memory. For example, for pilots to learn to fly, thousands of hours of practice are needed so that they develop the reflexes that enable them to act on instinct in life-and-death combat situations. Not all of these hours can be accomplished through actual flight time because of the risks associated with flying and the resources required to send aircraft out on a training mission. Flight simulators, which are designed to emulate every detail of an aircraft and its performance, offer a way to increase training frequency and duration without the costs and risks associated with real-world flight.

Today in cyberspace operations, most hands-on technical training occurs in lab environments with real or virtual hardware and software tuned to specific training objectives without regard for the holistic operating environment (i.e., the configuration of people, processes, and technologies that make up the cyber terrain, including command and control and intelligence functions). While the use of specific programs and commands may translate from the lab into procedural memory useful in the real-world, many variables change from the classroom to the "battlefield." Introducing a common human-machine interface that employs game elements and game design to facilitate learning and efficient operation may open opportunities for the cyber equivalent to the flight simulator. Learning may be further facilitated through the use of gamified motivation techniques, such as points, badges, and leaderboards. Training in this manner may encode in procedural memory the locations and processes in software that soldiers need in order to access relevant observable artifacts and therefore decrease T_{OODA}. As many practiced players of various roles operate tools and interact more efficiently through the game-like interface, fog of war may also decrease.

Orientation

In today's cyber operations environment, orientation often requires the assimilation of information from diverse sources, distributed via multiple methods and modalities that are often nonstandard. Oftentimes, this information works its way through intermediaries that induce loss to the original information. Once real-world operators or analysts have collected and fused actionable information, it often takes hours or days to orient to the information, decide a course of action, and finally enact that course of action.

Compare the above notional T_{OODA} of real-world cyber security operations with that of the real-time strategy game *StarCraft II*®. In *StarCraft*, a casual player can sustain a productivity level of 50 complex, meaningful, and multidisciplinary actions per minute (APM) while a proficient player can sustain 300 or more APM [7]. These numbers, while unlikely in real-world operations, represent the T_{OODA} speeds humans are capable of when presented with near-lossless interfaces to accurate information, capabilities, and real-time feedback. Developing an equivalent gamified interface to real-world operations may enable players to quickly observe the artifacts presented, orient to them with computational augmentation and automation, decide courses of action based on probabilities of effectiveness, and from within the same interface take actions or issue orders and guidance for others to take action. Such a game-like interface may decrease time and signal loss from sensor to decision maker and from decision maker to actuator, thereby decreasing T_{OODA} and fog of war.

Deep Insight

While serious games tend to capture structured data regarding the impact that observable artifacts, tools, and interactions have on metrics like fog of war and T_{OODA}, these data largely go uncaptured in today's real-world operational environment. In a common gamified platform, metadata associated with each of the OODA steps can be collected and used as immediate player feedback in the form of achievement badges, experience points, and ranking on leaderboards. These metadata can also be used for analytical inquiry into the efficacy of plans developed and tactics employed in real-world operations or the exercises that precede them.

Where to Begin

In order to apply game elements and game design techniques to military cyberspace operations' mission applications, such as battle management systems, we can leverage game design approaches, such as Hunicke et al.'s mechanics, dynamics, and aesthetics framework [8].

Mechanics describes the particular components of the game, at the level of data representation and algorithms. Dynamics describes the runtime behavior of the mechanics acting on the player inputs and each other's outputs over time. Aesthetics describes the desirable emotional responses evoked in the players when they interact with the game system [8].

MECHANICS AND GAME CONTENT

All games have rules, workflows, assets, levels, roles, and a variety of other mechanisms and content that enable gameplay. To understand these mechanics for the design of a gamified battle management system for cyberspace operations, we can turn to the Doctrine for the Armed Forces of the United States, which contains thousands of pages clearly defining, among other things, the intelligence, operations, and planning methodologies employed in all domains of conflict [9]. Applying this doctrine to the cyberspace domain requires us to research the specific functions and tasks described in cyber security guides and literature. By combining the discrete tasks necessary to secure and operate networks with the military concepts necessary to conduct full-spectrum military operations, we can define the mechanics of cyberspace operations. We have already started work to describe these mechanics and expect the results to feed future prototyping efforts for a gamified battle management system.

DYNAMICS

To keep players interested, game designers often create game elements such as time pressure or tension within the storyline of the game. However, these elements already exist in real-world military conflicts. While cyberspace operations are likely to have their dynamics driven by geopolitics or current in-contact operations, we must strive to understand these and other dynamic components as we gamify the cyber operations environment. We may want to put aesthetic mechanisms in place to convey dynamics; for example, we could add countdown clocks to indicate deadlines for countermeasure deployment or audio feedback to indicate success.

AESTHETICS

To look through the eyes of the player, we must consider the aesthetics of the game and the motivations (extrinsic or intrinsic) that drive them to play the game. While in traditional military system designs aesthetics are rarely considered, they are critical in the cyberspace domain. Because of the complexity of the cyber environment, potential players will always look for ways to decrease complexity, using the path of least resistance even if doing so inadvertently increases fog of war and T_{OODA}. Designing a user interface that considers how the interface will impact the user's mental and emotional state, that is intuitive to operate, and that is even fun to use may promote the gamified system's use over more familiar systems that do not consider the game mechanics necessary to decrease fog of war and T_{OODA}. The gamified mission application should at minimum provide users a venue that makes their role easier, more effective, and more motivating than do current methods and modalities, such as email and document-based approaches.

Summary

Lincoln Laboratory's Cyber Red/Blue game environment provides a repeatable methodology for measuring human behaviors that affect cyber security outcomes. Inclusion of real operational tools in the game environment will improve training and analysis results. With actual tools and the flexible Cyber Red/Blue measurement framework, it is possible to apply additional measurement qualities of mechanics, dynamics, and aesthetics to a gamified real-world environment that simultaneously measures and trains for the future. We look forward to developing this approach further. ■

References

1. C. von Clausewitz, *On War*. Berlin: DümmlersVerlag, 1832; 1874 translation by J.J. Graham available as a Project Gutenberg ebook at www.Gutenberg.org.
2. J.R. Boyd, "Patterns of Conflict," 1986 version edited by C. Richards and C. Spinney in 2007 for the Project on Government Oversight, Defense and National Interest, www.dnipogo.org/boyd/patterns_ppt.pdf.
3. National Cyber Range, Deputy Assistant Secretary of Defense for Developmental Test & Evaluation/Director, Test Resource Management Center website, https://www.acq.osd.mil/dte-trmc/ncr.html.
4. SANS Institute website, https://www.sans.org/.
5. DoD Cyber Awareness Challenge Training, https://ia.signal.army.mil/dodiaa/.
6. K. Robson, K. Plangger, J.H. Kietzmann, I. McCarthy, and L. Pitt, "Is It All a Game? Understanding the Principles of Gamification," *Business Horizons*, vol. 58, no. 4, 2015, pp. 411–420, http://dx.doi.org/10.1016/j.bushor.2015.03.006.
7. Y. Lejacq, "How Fast Is Fast? Some Pro Gamers Make 10 Moves per Second," NBC News, 24 Oct. 2013, http://www.nbcnews.com/technology/how-fast-fast-some-pro-gamers-make-10-moves-second-8C11422946.
8. R. Hunicke, M. LeBlanc, and R. Zubek, "MDA: A Formal Approach to Game Design and Game Research," *Proceedings of the Challenges in Games AI Workshop, 19th National Conference of Artificial Intelligence*, 2004.
9. U.S. Department of Defense, *Doctrine for the Armed Forces of the United States. Washington*, D.C.: CreateSpace Independent Publishing Platform, 2013.

About the Authors

Nancy L. Crabtree has more than 13 years of experience at Lincoln Laboratory, working in the area of cyber systems engineering and architecture for both the Information Services Department and the Cyber Security and Information Sciences Division. She has applied her work in cyber security and resilience to diverse Department of Defense (DoD) domains, including installation energy, next-generation radar, space systems, command and control, acquisitions, and critical enterprise infrastructure. Prior to joining the Laboratory, she worked in software development, implementation, and quality assurance for data communications, Internet, and telephony security companies, such as GTE Internetworking, BBN, and Boston Technology. She is an early member of the regional cross-sector Advanced Cyber Security Center and a member of the Military Operations Research Society, IEEE, and several DoD information-sharing groups. She holds bachelor's and master's degrees in technical business management and organizational behavior from Thomas Edison State University and has a Certified Information Systems Security Professional certification from (ISC)^2, an international cyber security organization.

Joshua A. Orr was an application developer and cyber operations analyst in Lincoln Laboratory's Cyber Systems and Operations Group. He was assigned to the field site at Fort Meade, Maryland. In his primary role as the senior technical advisor for the U.S. Cyber Command (USCYBERCOM) Capability Development Group, he was responsible for the analysis and design of solutions to a broad range of mission-critical technology requirements for national strategic operations. He is currently the deputy director of cyber operations for the 2020 U.S. census. Prior to joining the Laboratory in 2014, he served 14 years in the U.S. Air Force in the cyber and network engineering field, supporting intelligence and cyber operations missions. He has extensive experience in cyber operations and was responsible for designing and building the Cyber Command and Control Portal for Operations while assigned to USCYBERCOM. He holds several industry certifications, including GIAC Certified Incident Handler, GIAC Penetration Tester, and Security +. He earned his bachelor's degree in business administration at Grantham University in 2011 and is currently pursuing a bachelor's degree in computer science through the University of Maryland University College.