Technical Report 1408

Development of an Interactive Systems Thinking Assessment

Cory Adis
Personnel Decisions Research Institutes

Alexander P. Wind
U.S. Army Research Institute

Michelle Wisecarver
Katie Guarino
Jaclyn Martin
Lia Engelsted
Chelsey Byrd
Personnel Decisions Research Institutes

Phillip Mangos
Adaptive Immersion Technologies

April 2022

United States Army Research Institute
for the Behavioral and Social Sciences

U.S. Army Research Institute
for the Behavioral and Social Sciences

Department of the Army
Deputy Chief of Staff, G1

Authorized and approved:

MICHELLE L. ZBYLUT, Ph.D.
Director

Research accomplished under contract
for the Department of the Army by

Personnel Decisions Research Institutes

Technical review by

Peter Legree, U. S. Army Research Institute
Colin Omori, U. S. Army Research Institute

DISPOSITION

This Technical Report has been submitted to the
Defense Information Technical Center (DTIC).

REPORT DOCUMENTATION PAGE

1. REPORT DATE *(DD-MM-YYYY)*
April 2022

2. REPORT TYPE
Interim

3. DATES COVERED *(From - To)*
August 2018 – August 2020

4. TITLE AND SUBTITLE
Development of an Interactive Systems Thinking Assessment

5a. CONTRACT NUMBER
W911NF-18-C-0052

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER
622785

5d. PROJECT NUMBER
A790

5e. TASK NUMBER
1008

5f. WORK UNIT NUMBER

6. AUTHORS
Adis, Cory; Wind, Alexander, P.; Wisecarver, Michelle; Guarino, Katie; Martin, Jaclyn; Engelsted, Lia; Byrd, Chelsey; Mangos, Phillip

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Personnel Decisions Research Institutes,
1911 N. Fort Myer Drive, Suite 410
Arlington, VA 22209

8. PERFORMING ORGANIZATION REPORT NUMBER
N/A

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U. S. Army Research Institute
for the Behavioral & Social Sciences
6000 6TH Street (Bldg. 1464 / Mail Stop 5610)
Fort Belvoir, VA 22060-5610

10. SPONSOR/MONITOR'S ACRONYM(S)
ARI

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
Technical Report 1408

12. DISTRIBUTION/AVAILABILITY STATEMENT: Distribution Statement A: Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES
ARI Research POC: Colin L. Omori, Selection and Assignment Research Unit

14. ABSTRACT

Numerous jobs in the U.S. Army require Soldiers to work with or within systems, yet systems thinking is a talent that is challenging to measure accurately. Given the pervasiveness of systems across Army jobs and requirements, particularly in cyber-related occupations, the U.S. Army Research Institute for the Behavioral and Social Sciences (ARI) identified a need for an ecologically valid assessment tool that can identify personnel who have an aptitude for systems thinking. A first-person computer simulation with 3D graphics was developed as a vehicle to measure 13 characteristics related to systems thinking. Multiple assessment methods are employed to assess the various dimensions while the test taker engages in the simulation. The report describes each measure in further detail and provides information regarding the premise and objectives. Next steps for development and validation are discussed.

15. SUBJECT TERMS
Cognitive attributes, Noncognitive attributes, Creativity, Curiosity, Openness to information, Simulation, Systems thinking, Assessments

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
Unclassified	Unclassified	Unclassified	Unlimited Unclassified	50	Tonia S. Heffner
					19b. TELEPHONE NUMBER 703-545-4408

Technical Report 1408

Development of an Interactive Systems Thinking Assessment

Cory Adis
Personnel Decisions Research Institutes

Alexander P. Wind
U.S. Army Research Institute

Michelle Wisecarver
Katie Guarino
Jaclyn Martin
Lia Engelsted
Chelsey Byrd
Personnel Decisions Research Institutes

Phillip Mangos
Adaptive Immersion Technologies

Selection and Assignment Research Unit
Tonia S. Heffner, Chief

April 2022

ACKNOWLEDGEMENTS

We would like to thank personnel who contributed to this work through helpful comments and feedback after numerous hours spent on reviews and piloting activities, particularly Alisha Ness for her suggestions and feedback on early versions of the assessment. The authors would also like to thank Anna Guidubaldi, Tom Motz, Sidney Quaid, Austin Kotlus, and Gary Huffman for their tireless work to develop the functional software for the assessment.

DEVELOPMENT OF AN INTERACTIVE SYSTEMS THINKING ASSESSMENT

EXECUTIVE SUMMARY

Research Requirement:

Numerous jobs in the U.S. Army require Soldiers to work with or within systems, yet systems thinking is a talent that is challenging to measure accurately. Given the pervasiveness of systems across Army jobs and requirements, particularly in cyber-related occupations, the U.S. Army Research Institute for the Behavioral and Social Sciences (ARI) identified a need for an ecologically valid assessment tool that can identify personnel who have an aptitude for systems thinking. Previous research (Adis et al., 2017) identified multiple antecedent attributes and behavioral dimensions that are conceptually relevant to systems thinking. Advances in technology, such as game-based assessments, are providing new avenues to measure more complex constructs like systems thinking and identifying Soldiers who have an aptitude for systems thinking. Systems thinking involves understanding the activity and changing characteristics of a system as it operates and anticipating effects of actions upon the system to attain a particular outcome (Canali et al., 2017). Better assessment of systems thinking could enhance talent management at initial entry when military occupational specialties (MOS) or branches are determined or when making in-service assignments.

Approach:

A first-person video simulation with 3D graphics was developed as a vehicle to measure eight antecedent attributes and five behavioral dimensions related to systems thinking. The simulation has an alien invasion theme, in which the Earth is being threatened by an alien spaceship and the test taker is one of many individuals abducted and taken prisoner. Using an alien invasion theme is functional because the systems included do not need to be accurate representations of existing systems, and it limits test takers from having an advantage by already being familiar with the specific systems being used for the assessment. Multiple assessment methods are employed throughout the experience to assess the various attributes and dimensions while the test taker engages in the simulation. These methods include a memory test, situational judgment tests, a continuous performance test, rating tasks, multiple-choice tests, logical reasoning tests, perceptual tests, and creative problem-solving tasks.

Findings:

An initial functional version of the assessment was developed to measure four dimensions of systems thinking: Awareness of Elements, Identifying Relationships, Understanding System Dynamics, and Application to Mission. Test takers first explored systems in their environment, then recorded their view of the important elements comprising those systems (Awareness of Elements) and the relationships among those elements (Identifying Relationships). They then demonstrated their understanding through questions about feedback loops within those systems (Understanding System Dynamics) and applied that understanding through questions about the importance of the systems for accomplishing mission objectives (Application to Mission).

Measures of three antecedents of systems thinking were also developed. Two of these attributes, creativity and curiosity, were measured through behavioral actions collected as the test taker explores the ship. Each of these components were tested within dedicated interfaces where confounds can be reduced or mitigated. The other attribute, openness to information, required the test taker to answer questions that test the extent to which they noticed and retained information relevant to the overall objectives. Stimuli for this measure were presented throughout the course of exploration and are tested with a rating task at the end.

Finally, five antecedent cognitive attributes were incorporated using previously developed assessments that required various cognitive actions or judgments: hierarchical working memory, cognitive flexibility, cognitive complexity, pattern recognition, and spatial ability. As part of the premise of the game, the test takers had to complete these more traditionally-designed measures.

This report describes each measure in further detail and provides additional information regarding the premise and objectives. Areas of required development and testing are discussed.

Utilization and Dissemination of Findings:

Now that the assessment is developed, construct and criterion-related validity evidence will need to be collected and evaluated. Initial validation evidence is required for honing the measures and calibrating the test system material. The items that need to be evaluated for reliability and content can be added or removed to improve psychometric performance. Additionally, testing across multiple MOS is required to determine which measures to include for a variety of job contexts. Once validated, the battery could be used to identify Soldiers who have a higher likelihood of success in jobs that require systems thinking.

A description of the assessment context and measures were provided to members of the Manpower Accession Policy Working Group (MAPWG) during a regularly scheduled MAPWG meeting.

CONTENTS

LIST OF TABLES

LIST OF FIGURES

DEVELOPMENT OF AN INTERACTIVE SYSTEMS THINKING ASSESSMENT

Introduction

Systems are pervasive across Army jobs and requirements. Numerous jobs in the Army require Soldiers to work with or within systems, such as complex cyber systems, social networks, and the organizational systems officers lead and manage (Allen et al., 2009; Dedmond, 2014; Dennis, 2010; Pretz et al., 2003). The U.S. Army Research Institute for the Behavioral and Social Sciences (ARI) is conducting innovative research to develop and validate a means to identify personnel with an aptitude for systems thinking. An assessment to quantify an individual's systems thinking aptitude will enhance talent management selection and assignment decisions.

A system can be defined as a set of interconnected elements that is organized coherently and achieves a function or purpose (Meadows, 2008). Since the middle of the last century, several different systems thinking approaches or paradigms have been developed across different disciplines (Flood & Jackson, 1991; Jackson & Keyes, 1984; Midgley, 1997; Mingers, 2006). Each paradigm was developed for a particular situation and uses a specific approach and conceptualization of "systems." These paradigms look at systems from the perspective of different types of metaphors: systems as aggregations, organisms, feedback loops, thinking perspectives, and as a set of cognitive tools (Adis et al., 2013). Systems thinking ability involves understanding the activity and changing characteristics of a system as it operates and anticipating effects of actions upon the system to attain a particular outcome (Canali et al., 2017). Accordingly, those who have a tendency to engage in systems thinking are more likely to make decisions that enhance the well-being of the systems with which they interact (Thibodeau et al., 2016).

The importance of thinking in terms of systems has long been recognized across various scientific and applied domains, from engineering and biology to organizational processes (for reviews, see Cabrera, 2006; Midgley, 1997, 2003; Mingers, 2006). Despite the widespread conceptual use of systems thinking, few attempts have been made to measure systems thinking in individuals. Most measurement attempts are better suited for use in training and development rather than selection or classification due to reliance on self-report measures or scoring protocols that require human raters and are therefore costly and time-consuming to administer. For example, existing measures include self-reported systems thinking capacity (Randle & Stroink, 2012), system-based scenarios and a think-aloud protocol (Sweeney & Sterman, 2000, 2007), a game-based simulation followed by a short writing assignment (Shute, 2011; Shute et al., 2010, 2016), and the creation of causal loop diagrams (Shute, 2011; Shute et al., 2010, 2016).

In response to the lack of systems thinking assessments, ARI developed and validated a set of measures for five antecedents of systems thinking: (a) hierarchical working memory capacity (HWMC), (b) spatial ability, (c) cognitive flexibility, (d) pattern recognition, and (e) cognitive complexity (see Adis et al. 2017). Results generally provided support for the construct validity of these measures, although the level of support varied.

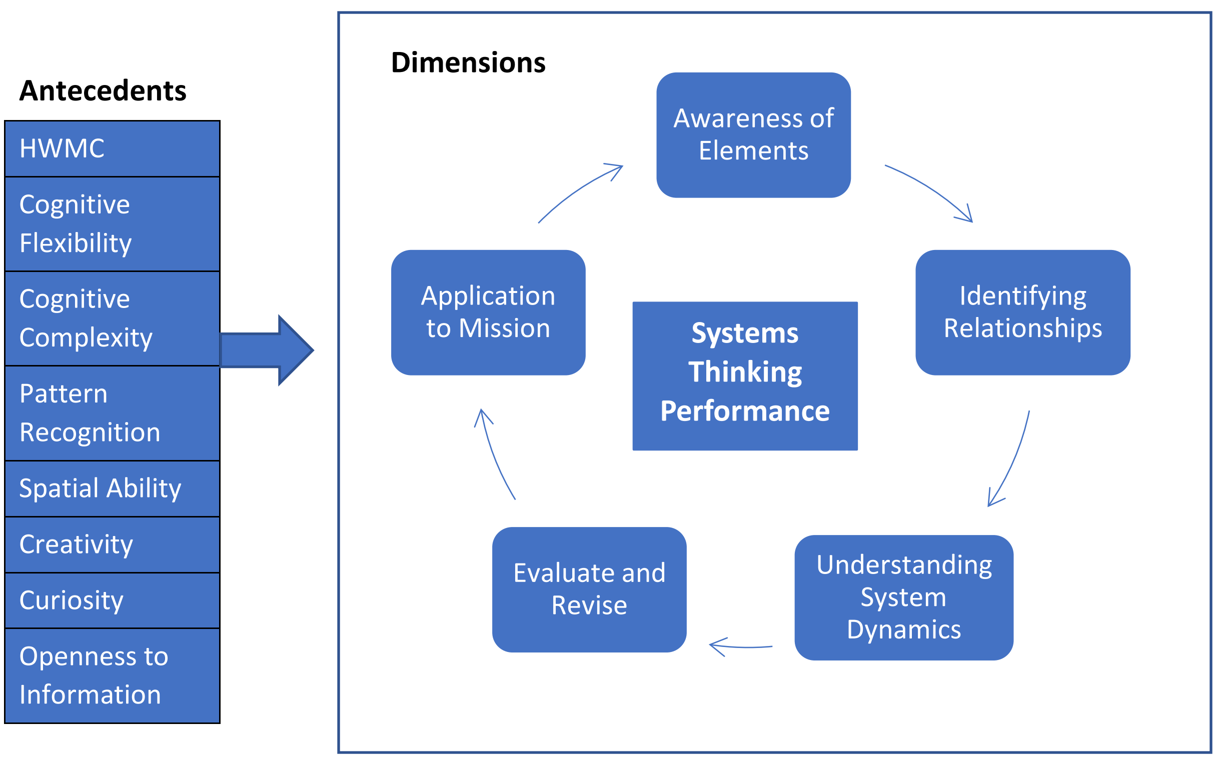
The current effort extends ARI's research by developing measures for three additional systems thinking antecedents (curiosity, creativity, and openness to information) and the five dimensions of systems thinking (Awareness of Elements, Identifying Relationships, Understanding System Dynamics, Evaluate and Revise, Application to Mission). The content for one of these five dimensions - Evaluate and Revise - was developed but not yet implemented in the assessment game. Further, this effort details steps taken to refine previously developed measures of the five antecedent cognitive ability constructs. Subsequent validation work will determine which measures from this entire set should ultimately be included in the final operational tool. In the following sections, a measurement approach for each dimension and antecedent construct, and the context of the first-person simulation are described. Finally, next steps for development and empirical validation are provided.

Describing Antecedents and Dimensions of Systems Thinking

Systems thinking requires a constellation of closely related abilities that enable individuals to (a) identify the elements of a system, (b) understand system relationships, (c) evaluate and revise system models, and (d) apply an integrated understanding of the system to a problem. Adis et al. (2017) developed a model of systems thinking, which has two components – a set of dimensions, and a set of antecedents (see Figure 1).

Figure 1

Model of Systems Thinking



The dimensions component suggests that systems thinking can be measured by examining five dimensions that mirror stages of a generalized, systems thinking problem-solving cycle. Each of these stages reflects aspects of problem solving that are critical for successful systems thinking performance. These dimensions are founded in the behaviorally based definition provided for systems thinking. The five systems thinking dimensions are defined in Table 1.

Table 1

Description of Five Dimensions of System Thinking

Dimension	Definition
Awareness of Elements	The ability to identify and describe components or subsystems of a problem situation.
Identifying Relationships	The ability to accurately identify structural or causal patterns among elements of a system or between a system and its environment.
Understanding System Dynamics	The ability to recognize feedback loops resulting in system behavior over time.
Evaluate and Revise	The ability to recognize and systematically test model assumptions in a proactive manner; the ability to accurately interpret data and their implications for system model accuracy.
Application to Mission	The ability to relate modeled system expectations back to the situation-at-hand and draw insights from system analogies.

The antecedent component of Figure 1 reflects a measurement approach that focuses on a set of distal antecedents. Systems thinking antecedents are enabling abilities that are necessary to effectively engage in the systems thinking process. Since systems thinking is a multi-stage process that individuals engage in constantly for a prolonged time, there are many other factors that will impact an individual's performance over that time. Given this, these enabling abilities are expected to show weak-to-moderate relationships with systems thinking due to the mediating or moderating influences of other constructs. The eight systems thinking antecedents are defined in Table 2.

Given the novelty of this type of systems thinking assessment, model antecedents and dimensions were included in the initial test battery to enable exploration of the relationships among the constructs. While empirical validation is needed, both are expected to predict systems thinking on the job. Some overlap is expected between components of the model, with antecedents predicting systems thinking performance directly as well as through relationships mediated by the dimensions. It is also expected that there may be overlap among the antecedents, among the dimensions, and/or between the two components. Nevertheless, efforts were taken to uniquely distinguish each construct through their operationalizations. Empirical evidence will be needed to delineate the relationships in the model and the role of each construct in predicting systems thinking performance.

Table 2

Description of Eight Antecedents to Systems Thinking

Construct	Definition
Hierarchical working memory capacity	Individual differences in the cognitive system responsible for storing information required for ongoing cognitive operations.
Spatial ability	The set of cognitive abilities that allow an individual to process visual stimuli among distractors, understand a visual scene, and faithfully encode and mentally manipulate visual objects or spatial relationships.
Cognitive flexibility	The readiness with which a person's concept system changes appropriately in response to environmental stimuli.
Pattern recognition	The ability to find repetitions in sequences of objects or data, or in the rules governing their sequence. The ability to detect deviations in normal patterns of behavior.
Cognitive complexity	The extent to which an individual is able to differentiate an event into distinct elements and integrate or make connections between those elements.
Curiosity	Possessing the desire to know or learn about people, things, and relationships and taking action to investigate them.
Openness to information	Taking note of new information or points of view that are relevant to models or plans and considering ways in which the information impacts those models or plans.
Creativity	Producing ideas or products that are novel or unusual and provide a useful approach or perspective in a given situation.

Dimensions of Systems Thinking

Due to the lack of empirical research on measurement, an analytical approach was used to identify dimensions required in the systems thinking process. The five systems thinking dimensions identified by Adis et al. (2017) are rooted in five themes that emerged across various systems thinking paradigms found in the literature: (a) identifying the elements of the system; (b) understanding dynamic relationships within the system; (c) determining and modeling how the elements are related; (d) engaging in a process of evaluating and revising one's understanding of the system, and (e) applying an integrated understanding of the system to the problem (for details see Adis et al., 2017). The five themes also align with elements of the Military Decision-Making Process (MDMP) and problem solving cycles (Pretz et al. 2003, pp. 3–4). The similarities between systems thinking and decision making or problem solving suggest that a cognitive problem solving framework is a useful theoretical foundation for understanding the systems thinking process, particularly in ill-defined situations, which present distinct challenges in identifying and representing problems in order to develop a solution strategy (e.g., Pretz et al., 2003). Elements of both systems thinking and problem solving reflect a cyclical process that moves from an initial problem state toward a goal state and requires representing relationships between their elements and monitoring the process as it progresses. In a sense, systems thinking can be considered a unique form of problem solving in which the problem requires building an understanding of the relationships and dynamics that exist within the system. These dimensions are described below.

Awareness of Elements

Becoming aware of a system's elements is often the first stage in the systems thinking process. This entails identifying important components of the system, as well as the system's boundaries. Relating systems thinking to a problem-solving process, this stage is equivalent to defining the problem space. Individuals who are better able to define the key elements in the problem space should be better able to develop solutions. These individuals would be more likely to recognize leverage points for affecting the system through their understanding of system elements. The solutions generated by those high in Awareness of Elements are likely to consider more second- and third-order effects than the solutions of those individuals who are unable to recognize system elements. Being able to depict system elements provides the foundation for being able to identify and understand system relationships and dynamics – other critical dimensions of systems thinking. In a sense, Awareness of Elements is a gateway to other systems thinking processes, so individuals high in this construct would also be likely to engage in more systems thinking in general.

Identifying Relationships

Dissecting relationships among system elements is an important second step after the system elements are identified. This dimension allows a systems thinker to define dyadic relationships that describe how elements work together, interact, or cause some type of response in one another. Recognizing these relationships is a precursor to understanding more complex interactions and dynamics in the system. Individuals who are able to thoroughly identify all relationships among recognized elements would be better able to generate a mental model of the system, and likely to be more accurate when mentally simulating interactions in the system. As a result, these individuals will more easily recognize second- and third-order effects, leading to decisions that better account for these consequences.

Understanding System Dynamics

Once simple dyadic relationships are discerned, the systems thinker has the potential to dig deeper into the interrelatedness of the system elements. Understanding System Dynamics allows a systems thinker to understand the changing characteristics of a system as it operates and to anticipate changes in system behavior based on the effects of actions upon the system. This involves an understanding of system characteristics, such as feedback loops that go beyond cause-effect relationships and create cyclical patterns. Individuals who are better able to understand dynamic relationships will be able to more quickly recognize systems. With training, these individuals will be able to recognize when to engage in systems thinking and what techniques to use.

Evaluate and Revise

After a systems thinker creates a mental model of the system, they must develop a deeper understanding of a problem situation by reconciling a simplified model with real-world, operational data. The Evaluate and Revise dimension captures the ability to revise their understanding of a system by considering new information, some of which may contradict prior assumptions. Individuals who look for and more quickly adapt to new information will be better suited for fast-paced, dynamic environments because these individuals will be able to keep their models current. By frequently checking for deviance between models and reality, people who score highly on the Evaluate and Revise dimension will plan more effectively and minimize the scope of surprises.

Application to Mission

The last stage of the systems thinking process entails applying an integrated understanding of the system to inform a decision or solve a problem. By developing an understanding of the system, the systems thinker can identify leverage points for affecting change in the system. Individuals who score highly on the Application to Mission dimension will be able to recognize how leverage points can be used to achieve proximal goals while keeping the overarching mission in mind. Additionally, individuals high in Application to Mission remain grounded during system exploration, keeping their learning, modeling, and experimentation in check by consistently thinking back to the mission and the constraints imposed by the problem or decision at hand.

Antecedents of Systems Thinking

The previous section discussed five dimensions linked to key stages of systems thinking performance. In addition to the more proximal dimensions, Adis et al. (2017) proposed that there is a set of distal individual difference constructs that underlie effective performance of systems thinking (). These constructs both enable the five dimensions and are themselves potentially important predictors of an individual's systems thinking performance.

Eight distal constructs were identified to measure underlying abilities that are basic ingredients for systems thinking (Adis et al., 2017). Of these eight abilities, Adis et al. developed and tested five constructs: hierarchical working memory capacity, spatial ability, pattern recognition, cognitive flexibility, and cognitive complexity. Assessments for the three remaining constructs were developed in the current effort. All eight constructs are listed in Table 2 and the relationship of each construct to systems thinking is detailed in the following paragraphs.

Hierarchical Working Memory Capacity

Working memory capacity has been identified as relevant to successful problem solving (e.g., Hambrick & Engle, 2003). Given that systems thinking involves building, testing, and monitoring models that represent systems and subsystems nested within those systems, working memory capacity that spans multiple levels - hierarchical working memory - is particularly relevant. Hierarchical working memory capacity can contribute to the tasks of building the mental and/or graphical representations of the systems, and being able to represent problems mentally is a key element of problem solving (Pretz et al., 2003), which is central to systems thinking.

Spatial Ability

Spatial ability is proposed to facilitate the processes of building, testing, and monitoring models of systems. A common theme across systems thinking literatures is the ability to "see" the structure of systems, which may reflect this underlying aspect of systems thinking ability. Spatial ability should contribute to a systems thinker's ability to process mental and physical models of systems, as well as graphs or diagrams of system behavior. Given the complexity of systems thinking and reliance on models to understand and describe these complexities, spatial ability is a logical antecedent to systems thinking.

Cognitive Flexibility

Cognitive flexibility is a potential precursor to the mental shifts that are required for the perspective taking aspects of systems thinking, in which an individual must alternate between a micro-level perspective and a macro-level perspective in order to understand a system. Cognitive flexibility enables an individual to entertain incongruous or conflicting pieces of information and is conceptually related to other aspects of systems thinking not included in this research, such as model testing, nonlinear thinking, thinking across time, and critical thinking.

Pattern Recognition

Pattern recognition is an important underlying ability for many of the cognitive processes essential for systems thinking, such as recognizing systems emergence and identifying system structure. This capability facilitates the identification of structures and patterns within a system, as well as identifying anomalies or elements that do *not* fit within the system or expectations for system behavior. This could be particularly relevant when operating to protect systems where anomaly detection can identify potential problems or threats to the system.

Cognitive Complexity

Cognitive complexity may underlie skills such as recognizing systems emergence, identifying structures, critical thinking, and metacognition, which have been identified as key systems thinking skills (Adis et al., 2017). An individual who is better able to see distinctions and interrelationships is likely to be better at systems thinking (Cabrera et al., 2008).

Curiosity

In the systems thinking context, curiosity drives deep investigations of the problem space, both as a means of problem solving and as a way of enhancing awareness of potential problem situations. Curiosity encourages an individual to ask the questions that are needed in order to identify system elements and their relationships, and to always be watching for new and changing information.

Openness to Information

Openness to information is important to ensure that new information that arises is considered and integrated into the understanding of a system. Without openness to information, individuals may develop mental models of situations based on immediately available information, then resist integrating new and emerging information into their models. Thus, this construct is likely to underlie an individual's ability to Evaluate and Revise their understanding of the system.

Creativity

The impact of creativity on systems thinking is likely broader and more indirect than some of the other individual characteristics being measured. Creativity affects systems thinking by enabling individuals to approach various aspects of the systems thinking process in novel ways. For example, enabling them to take unique perspectives in identifying elements of a system, have unusual ways of understanding how the elements are related, thinking of innovative ways to model and test system elements, and applying systems to understanding a problem in different ways.

Relationship between Antecedents and General Intelligence (g)

Since working memory capacity is highly *g*-loaded (Conway et al., 2003; Engle et al., 1999), Hierarchical working memory capacity is also expected to load heavily on *g*. Given the unique formulation of this working memory task, additional research is needed to determine the degree of overlap with *g* and the utilization strategy for this measure.

Spatial ability has been found to be independent of general intelligence (e.g., Rimfield et al., 2017) and therefore, the measures of spatial ability in this battery are expected to have low correlations with *g*. This property makes spatial ability advantageous as part of the battery or as a stand-alone predictor.

Cognitive flexibility is expected to have a moderate relationship with *g*. One operationalization of cognitive flexibility developed in this project aligns with the mental set switching executive function (Miyake et al., 2000) and is therefore likely to show high *g* loading. The other operationalization of cognitive flexibility is more similar to the personality dimension or behavioral tendency than an executive function and is therefore likely to show lower *g* loading. Empirical evidence will assist in determining the degree of overlap between the two cognitive flexibility tasks.

Pattern recognition is often related to general intelligence and some operationalizations are even used to measure fluid intelligence. The present operationalizations, however, require less induction of problem structure, which should reduce the similarity with *g*. Both the anomaly detection and multi-level figure recognition operationalizations are expected to have a low to moderate degree of overlap with *g*.

As with working memory, cognitive complexity is expected to be highly *g*-loaded. Given the unique operationalization developed for this project, performance on this task is also likely to be impacted by a number of other non-*g* factors. Depending on the results of validation testing, it may be determined that this construct does not adequately predict above measures of intelligence.

Curiosity is both a cognitive and a motivational personality construct. Some prior research has linked curiosity with achievement (e.g., Raine et al., 2002), suggesting a correlation between these constructs. However, other research (e.g., von Stumm et al., 2011) suggests that intellectual curiosity is related to intelligence, but a distinct predictor of achievement. As such, we expect moderate overlap with general intelligence, but some distinction as well.

Openness to information is similar to curiosity and cognitive flexibility in that it reflects a willingness to engage in cognition around new information. Though there is little existing research aimed at analyzing this construct, we expect that openness to information will have a similar relationship with intelligence as these two other constructs. As such, it is expected that openness to information will be moderately correlated with *g*.

The relationship between creativity and intelligence has been examined in psychology and education literatures for decades. Initial conceptualizations of these two constructs suggested that they are distinct and unrelated (e.g. Guilford, 1967; Wallach & Kogan, 1965). However, more recent studies, beginning with Carroll (1993) have concluded that creativity and intelligence are more similar than different (Kim, 2005; Silva, 2015).

Development of the Systems Thinking Assessment

In order to measure the five proposed dimensions of systems thinking and the eight antecedent abilities and attributes, we combined the dimensions and antecedents into one cohesive computer-based instrument – a systems thinking assessment. The systems thinking assessment can be conceptualized as a first-person puzzle adventure game. The test taker is prompted on where to go and what to investigate, though not on the specifics of what to do. These directive prompts are given to keep the test taker on course and making progress. The test taker navigates through the environment and interacts with the test systems by clicking locations on the screen. When the test taker has made the threshold number of observations about a test system, they are prompted to complete the test portion for that system. The test portion consists of a sequence of activities for Awareness of Elements, Identifying Relationships, and Understanding System Dynamics. Other construct tests are administered throughout the session (curiosity, creativity) or at the end of the narrative (openness to information, Application to Mission). The systems thinking assessment does not require any preexisting knowledge or systems thinking experience. It asks subjects to engage in some of the processes relevant to systems thinking during a simulated problem scenario.

The simulation has an alien invasion theme, in which the Earth is being threatened by an alien spaceship and the test taker is one of many individuals abducted and taken prisoner. Using an alien invasion theme is useful because (1) the systems used do not need to be accurate representations of existing systems, and (2) it prevents any test taker from having an advantage by already being familiar with the systems being used for the assessment.

In the scenario, while most prisoners have been put under the mind control of the aliens, the test taker remains mentally free and takes on the mission of collecting information about the aliens and their ship to pass back to forces on Earth. The test taker is introduced to a friendly alien who puts the player in contact with Earth forces and directs the player to a software tool that can help guide them to make their observations. The software tool (called the Automated Digital Assistant Module or ADAM) prompts the player to look for specific systems or to go to specific locations. ADAM also helps the player analyze the information they collect through the Mission Console interface that houses many of the assessments. The test taker observes and gathers information about systems in their environment and has to identify the patterns that emerge. The software simulates and provides information regarding multiple systems, which the test taker can explore and model. Once a system has been explored on the ship, the test taker goes to the Mission Console to record their observations, thereby completing the relevant assessments. The test taker is guided through entering observations for four systems, then concludes the session by submitting those observations, completing some mission-related questions, and boarding an escape vehicle for Earth.

In the following section, we first provide an overview of the different systems developed for the simulation context, then describe the measurement concepts for the five dimensions (Awareness of Elements, Identifying Relationships, Understanding System Dynamics, Evaluate and Revise, Application to Mission). Next, we describe revisions that were made to the five previously developed systems thinking antecedents (hierarchical working memory capacity, pattern recognition, cognitive complexity, cognitive flexibility, and spatial ability) based on the initial construct validity evidence reported in Adis et al., 2017. Finally, we describe the development of the remaining three systems thinking antecedent measures (curiosity, creativity, and openness to information).

Development of Test System Content

Multiple systems were explicitly developed to enable subjects to engage in the key processes relevant to the systems thinking dimensions. Test taker actions related to these dimensions are captured through the software's Mission Console, which prompts the test taker to investigate each of the ship's systems.

To build the assessment, several ship systems were designed, developed, and integrated into the storyline. A template was created to specify the system information, describe opportunities the test taker would have to observe and discover system concepts, and define the assessment material for each of the five systems thinking dimensions. The structure of the template can be seen in Appendix A. Initial concepts for systems led to the development of six system ideas based on adherence to two requirements. First, the test systems needed to include three or four subsystems so that the focus could shift to different aspects of a larger system and potentially reflect different system characteristics. Second, it was important for at least one subsystem to fit within common system patterns in systems dynamics, called system archetypes. The initial six system ideas can be seen in Table 3.

System archetypes are patterns of behavior that develop from a small set of building blocks: *feedback loops* (either reinforcing or balancing) and *delays* (Senge, 1990; Williams, 2007). These archetypes reflect emergent conditions resulting from persistent relationships over time (Senge, 1990). Senge (1990) lists nine examples of system archetypes that include different combinations and variations of feedback loops and delays. Identifying a system archetype allows one to gain a greater understanding of the relationship of the system elements and to identify what changes could impact the system's performance (Braun, 2002).

In systems with a *reinforcing feedback loop*, small changes amplify, or build upon themselves, and continue to either grow or decline when unchecked. For instance, if a supervisor sees potential in an employee, they may afford that employee more opportunities and resources to succeed. Unchecked, this self-fulfilling prophecy, or snowball effect, would continue such that the employee's success would increase, and they would continue to be afforded increasing opportunities and resources. Conversely, the opposite is also true such that an employee seen as low in potential is afforded fewer opportunities and resources, which results in what appears to be declining performance. This type of reinforcing feedback system aligns with the Army adage that "Units are good at things in which the commander is interested" (Williams, 2007, p. 19).

Table 3

Overview of Test Systems

Test System	Description	Built into Current Iteration
Sonic Elevator	The Sonic Elevators enable the ship's crew to travel between wedges and the levels of wedges on the ship. Test takers can learn how to access the elevators and how to use the elevator to travel between levels of one wedge. Eventually, test takers acquire the ability to select their destination floor and travel there automatically.	✓
Sentinel Guards	As guards, the Sentinels operate at two levels of activation: low and high. In low activation mode, the Sentinel stands in one place, occasionally making small movements, but remaining relatively stationary. Sensors are fully engaged, and most processing is done to recognize patterns in the physical environment. In high activation mode, the Sentinel moves deftly among the prisoners, expressing a large behavioral repertoire and action patterns.	✓
Energy Reactor	The Energy Reactors are the key energy source for each of the four lateral wedges of the ship. Energy from these reactors supplements the energy generated from the arboretum and is used to power various energy consuming projects.	✓
Swarmbot Defense System	The Swarmbots are a major defense system for the ship. They are used to guard the ship's exterior and screen all traffic that comes towards the ship. In the interior of the ship, the Swarmbots are used to move dangerous prisoners and restrict access to secure areas. Swarmbots have lasers they use to screen aliens and passengers for badges that signify that they are allowed to pass the Swarmbots.	✓
Water Shortage	Water on the ship is a closed system – no new water is added, and nearly all water that is used is reclaimed and recycled. Normally, recycling capacity far exceeds the demand for water, but due to numerous circumstances, the water system has come unbalanced and the water supply is dwindling. Typically, the use of the water is not controlled, but now that the water system is having trouble producing clean water quickly enough, the ship leaders have started to impose rations and restrictions.	X
Mind Control Technology	Mind control is initiated with a nasal implant that is controlled by wireless technology connected to a Hivemind. To continue the effects of the mind control after 24 hours, the prisoner must complete a series of training exercises to generate the electromagnetic pulses internally.	X

Systems with a *balancing feedback loop* seek to produce stability (Senge, 1990). That is, if an action is increasing, a balancing loop decreases it. If an action is decreasing, a balancing loop increases it. Ultimately, the balancing feedback is attempting to move the current state (the way things are) to a desired state (goal or objective). In a socio-political scenario that a Soldier might encounter while on deployment, there could be resistance to an implemented change that disrupts a country's norms, creating a balancing feedback loop whereby people desire to return to the traditional norms and ways of doing things. Finally, a *delay* is an interruption between an action and the consequence, or reaction that reduces the temporal proximity of the two and makes it more difficult to discern causality (Senge, 1990). Army systems frequently observe delays. For instance, there is often a time delay between submitting an operational needs statement for a new piece of equipment and actually having it fielded (Williams, 2007).

Each system was designed to contain an archetype subsystem as well as other subsystems designed to imitate systems of various types (e.g., mechanical systems, dynamical systems, complex adaptive systems, social systems) or subsystems that referred to a particular system characteristic (e.g., cybernetic control, nonlinearity, emergence). The concepts for these subsystems also had to integrate with the other related subsystems as well as the overarching storyline.

Of the six system ideas considered for this project, four were selected for implementation in the first version of the systems thinking assessment. The four selected systems are identified with a check in the third column of Table 3. During the assessment, the test taker is given instructions to explore each of the systems and learn how the system works and how it might be leveraged. After exploring the system, the test taker reports their understanding of the system through a question-and-answer format for each of the dimensions.

Each of the systems is relatively complex so has multiple subsystems identified. To make the identification of elements sufficiently challenging, test takers are asked to identify the key elements associated with specific subsystems; for example, identifying the key elements related to elevator entry and exit, not just identifying an element as relevant to the elevators. To make the task of identifying relationships sufficiently simple, test takers are asked only to identify direct relationships (a change in an element results directly in a change on another element) and not indirect relationships (a change in an element has effects on another element through an intervening element). A description of each of the systems and subsystems follows.

Sonic Elevators

The Sonic Elevators enable the ship's crew to travel between sectors of the ship and between levels within any given sector on the ship (see Figure 2). Test takers can learn how to access the elevators and how to use the elevator to travel between levels within one sector. Initially, they must travel with other crewmembers, then can try to pilot the elevator in manual mode. Once the test taker completes the Understanding System Dynamics test for the Sonic Elevator, the test taker can acquire a "drive" stick that enables them to select their destination floor without needing to "pilot" the elevator manually.

Figure 2

Screenshot of Sonic Elevator Doors

Key subsystems of the Sonic Elevators include Elevator Usage, Elevator Call Room Entry/Exit, Automatic Navigation, and Manual Navigation. Additional details follow regarding each subsystem.

- The Elevator Usage subsystem is an overarching system that includes key elements and relationships for accessing and operating the elevators. Many of the other subsystems are embodied by this subsystem but may also include elements that are not part of the Elevator Usage subsystem. The Elevator Usage subsystem currently includes 17 key elements for the Awareness of Elements module. This subsystem was not tested with the Identifying Relationships module since most of the relationships are included under other subsystems.

- The Elevator Call Room Entry/Exit subsystem entails a set of rules and procedures for accessing and leaving the elevator call rooms (an elevator loading and unloading area). This subsystem currently includes nine key elements and nine direct relationships among those key elements.

- The Automatic Navigation subsystem entails rules for how to activate the elevators' Automatic Mode, select a destination, and initiate the journey. This subsystem is first learned about vicariously, by watching other crewmembers using a console in the elevator call room to select a destination and activate the elevators. This subsystem currently consists of nine key elements and eight direct relationships.

- The Manual Navigation subsystem includes rules and controls for piloting the elevator manually and is explored after observing automatic mode, but before the test taker obtains the ability to travel in automatic mode without an escort. This subsystem currently includes 12 key elements. It also contains a feedback loop that matches a common system archetype and is, therefore, suitable for use in the Understanding System Dynamics module.

Sentinels

Sentinels can be found throughout the ship, but only cause trouble for test takers when they are in guard mode. An image of the sentinels can be seen in Figure 3. Sentinels in guard mode have orange or red lights on the front of their chests and on their helmets. In order to access some areas of the ship, the test taker must find a way to get past the orange and red guards. In addition to serving as guards, the Sentinels can also function as laborers or couriers; in these modes, they have blue and green lights on their suits, respectively. Sentinels in these modes do not interfere with the test taker, but the test taker must learn about the functionality of these modes.

Figure 3

Screenshot of Sentinels

Test takers must explore the overarching Sentinel robot (embodying system), as well as subsystems that describe the Sentinels' Modes and Uses, Mode Changes, and viable ways to circumvent the Sentinels in Guard Mode. Additional details follow regarding the subsystems.

- The embodying Sentinel system includes 14 key elements of the physical form of the Sentinels, including their structure, states, and routines. Like the Elevators, the overarching system was not tested in the Identifying Relationships module.

- The Modes and Uses subsystem is focused on the different modes in which the Sentinels can operate, and rules for how those different modes can be used. The reasons for these modes stem from the internal programming of the Sentinels, and their protocols must be deduced by observing Sentinel behavior. This subsystem includes 15 key elements and nine direct relationships.

- The Mode Changes subsystem deals with the logic behind conversion from one mode to another. It was not tested in the Awareness of Elements module because all elements for this subsystem are redundant with other subsystems. There were eight direct relationships in the Identifying Relationships module for the Mode Changes subsystem.

- The subsystem, Guard Mode, involves understanding how to get around the Sentinels in Guard Mode and consists of 12 key elements. The Guard Mode subsystem also includes a feedback loop archetype and was therefore used in the Understanding System Dynamics module.

Energy Reactor

The Energy Reactors are the key energy source for each of the four lateral sectors of the ship (see Figure 4). Energy from these reactors supplements the energy generated from an arboretum on the ship and is used to power various energy consuming activities. Test takers explore activating the reactors and using the reactors' energy to damage the alien ship. The subsystems are: the embodying system (the Reactor), Activating the Reactor, the Reaction Containment Mechanism, and Triggering the Failsafe Mechanism.

- The embodying Reactor is the subsystem describing the physical reactor and all its components. This subsystem currently consists of 28 components. As with the other test systems, the embodying system was not tested in the Identifying Relationships module.

- The subsystem for Activating the Reactor involves figuring out the sequence of settings required to turn on the Reactor. It currently consists of 10 key elements and six direct relationships.

- The subsystem for the Reaction Containment Mechanism deals with the mechanisms that kept the reaction from getting out of control when the reactor is operating normally. This subsystem currently consists of six key elements and six relationships. This subsystem also contains a common system archetype which is tested in the Understanding Systems Dynamics module.

- The subsystem for Triggering the Failsafe Mechanism centers on the conditions and events that are required for the Reactors' failsafe containment mechanisms to be activated. This subsystem currently consists of six key elements. It also contains a common system archetype included in the Understanding System Dynamics module.

Figure 4

Screenshot of Energy Reactors

Swarmbot Defense System

The Swarmbots are a major defense system for the ship (see Figure 5). They are used to guard the ship's exterior and screen all traffic that comes towards the ship. In the interior of the ship, the Swarmbots are used to move dangerous prisoners and restrict access to secure areas. Swarmbots have lasers they use to screen aliens and passengers for badges that signify that they are allowed to pass the Swarmbots. The Swarmbots are so named because of their swarm behavior. When Swarmbots aggregate together into a collective group, they follow particular rules of movement. Swarmbots can also aggregate into larger Swarmbot formations that look similar in shape regardless of size.

Figure 5

Screenshot of Swarmbots

There are five subsystems of the Swarmbots: the embodying subsystem, Swarmbot Activation Zones, Rules of Individual Behavior, Rules of Collective Behavior, and Internal/External Availability.

- The embodying subsystem for the Swarmbots includes all physical aspects of the bots themselves, as well as some of the protocols for their behavior and the technology that controls the location of their deployment.

- The subsystem for Swarmbot Activation Zones includes mechanisms and protocols for creating and regulating the Swarmbot zones of activity. These include internal and external defense areas, as well as regions within those defense areas where the Swarmbots ignore, monitor, or attack foreign objects. There are currently 12 key components of this subsystem and 11 direct relationships.

- The Rules of Individual Behavior subsystem is a system of rules for what Swarmbots seek out in their environment. The subsystem consists of five key elements and seven direct relationships.

- The Rules of Collective Behavior subsystem consists of the system of rules for how Swarmbots behave in groups. These rules give the Swarmbots their swarming behavior, as well as flocking behavior, and govern how and when smaller bots aggregate to larger formations. This subsystem currently consists of 8 key components and 11 direct relationships.

- The subsystem for Internal/External Availability describes the dynamic balance between the number of Swarmbot resources outside the ship versus the number of Swarmbot resources inside the ship. This subsystem consists of five key elements. It also contains a common system archetype included in the Understanding System Dynamics module.

Development of Systems Thinking Dimension Measures

The systems thinking dimensions are assessed using a variety of question and answer-based measures. The test taker is guided through each assessment as details about the ship's systems are discovered through their exploration. The test taker can review elements of the system in the Mission Console as they explore and develop a model of the system relationships. The Mission Console is an interface accessible to the test taker throughout the ship where the test taker receives assignments and is able to store information about the systems. It houses activities for each of the systems thinking dimensions. Once the test taker has collected all the required observations about a specific system, they are allowed to complete activities in the Mission Console regarding that system.

Test takers focus on one systems thinking dimension at a time for any given system, starting with Awareness of Elements. Once they believe they have identified all of the elements of a given system, they submit their concept of the elements in the Mission Console. After completing the Awareness of Elements module, the test taker moves to the Identifying Relationships module. In this module, the test taker selects pairs of elements from a subsystem and joins them with a verb or verb phrase to create causal sentences about the elements. Given a list of available elements and verb phrases, the test taker is asked to describe all the direct relationships among the elements.

After completing these first two modules for a given system, the test taker completes Understanding System Dynamics questions about that system. This component of the battery consists of multiple-choice questions where the test taker must identify the type of feedback loop described in a text-based stem or must select the most accurate graphic depiction of a feedback loop from among several response options. Very early in the software, test takers are given a brief tutorial on how to identify feedback loop types and basic concepts for how to model feedback loops graphically. They are then given general questions to test their understanding of this information. Later, after exploring the systems on the ship, test takers complete similarly structured questions pertaining to the feedback loops in the explored systems.

If the Evaluate and Revise module is implemented in the future, the test taker would next begin the Evaluate and Revise assessment. The test taker would be told that there may be new information available, and they are prompted to review a specific system location for each system to check for new information. For each system, the test taker is sent to a predetermined area of the ship and allowed to explore that area further. Additional pieces of information about the systems will be available then that were not available previously. The test taker will then have the ability to update their conceptualizations based on their review of the system. The content for the Evaluate and Revise module was designed and will be described in this report, but it has not yet been implemented in the assessment.

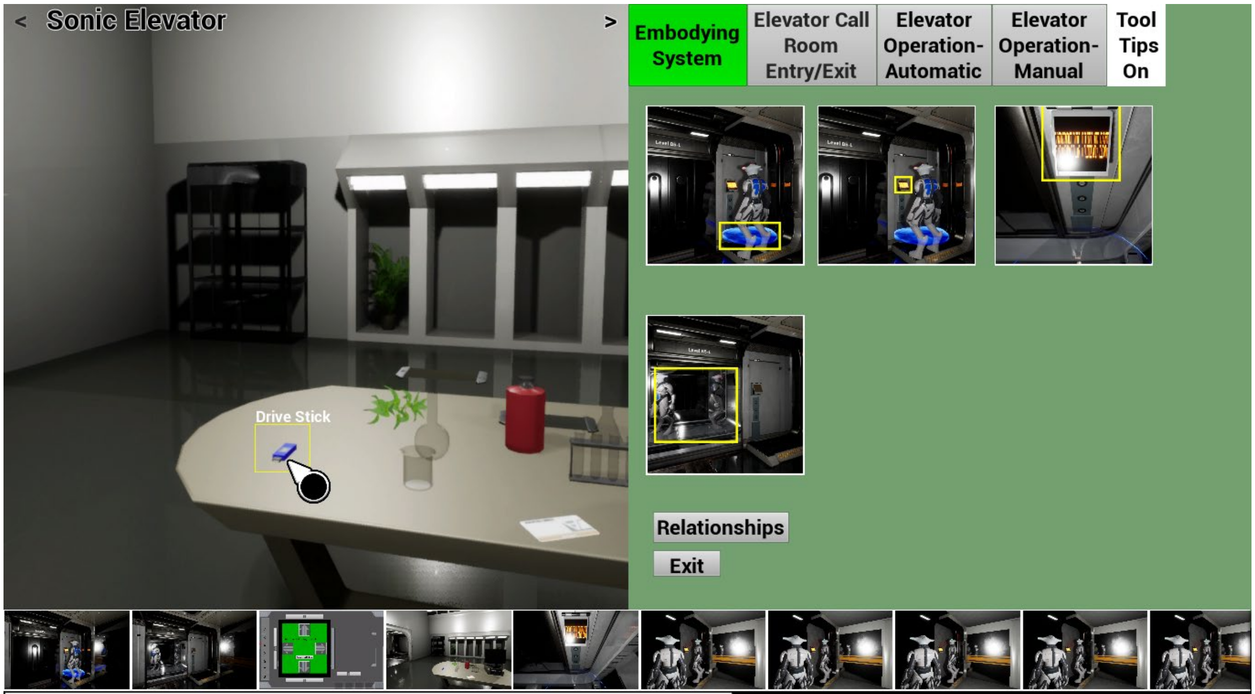
In the final module, Application to Mission, the test taker must demonstrate an understanding of the implications of system relationships and dynamics by answering multiple-choice questions. The systems thinking assessment calculates scores on the systems thinking dimensions for each test system and an overall score across the systems. Specific measurement approaches for each dimension are detailed below.

Awareness of Elements

In the Mission Console, the test taker is presented with a graphic list of their observations for a given system (see Figure 6). Each observation in the observation menu contains one or more elements that can be added to one or more subsystems. Some may be important elements, and others are distractors. The Awareness of Elements module assesses how many elements the test taker considers. By definition, there is not one most important element in a system, but there are elements that are involved and those that are not involved in a particular subsystem. Numerous raw data points will be collected such as the number of elements included for each subsystem, the number of correct or incorrect selections for each subsystem, or the total number of correct or incorrect elements across the subsystems and system. Since elements for a subsystem can be involved or not involved and the test taker can select or not select each element, a sensitivity metric may be the most appropriate calculation because sensitivity metrics take into consideration hits, misses, correct rejections, and false alarms. Additionally, a test taker's bias toward including or not including elements will be captured.

Figure 6

Awareness of Elements Module



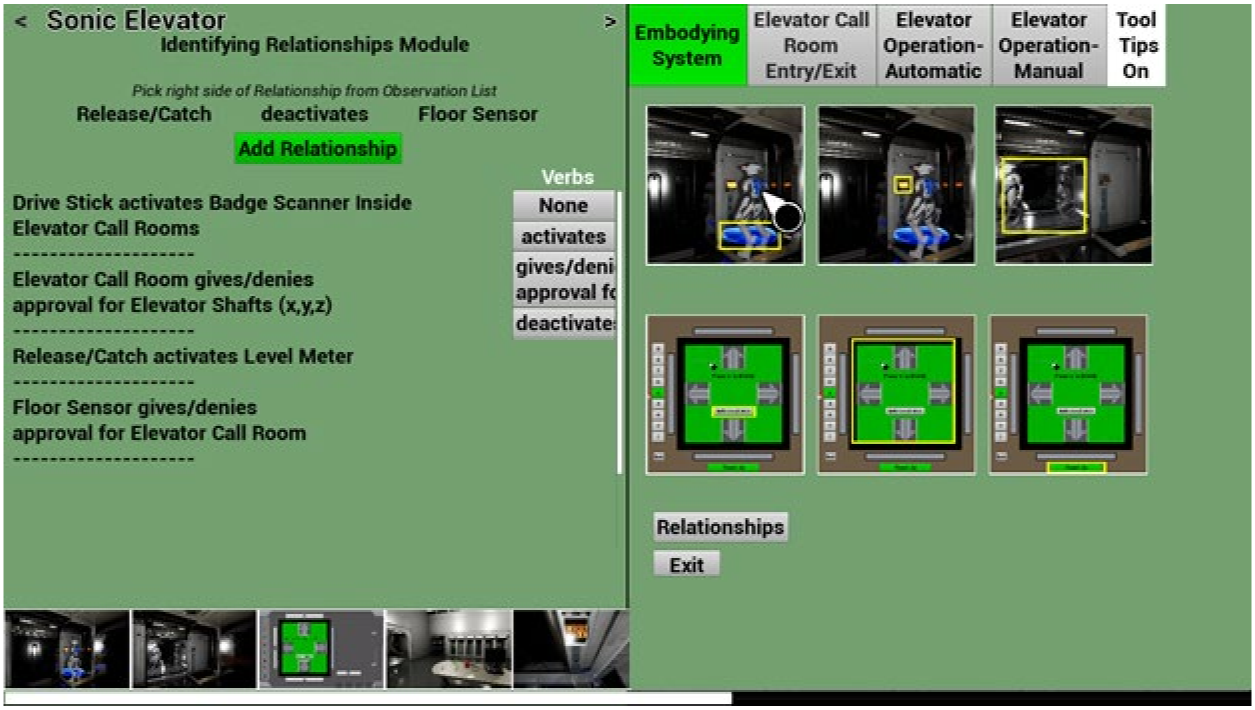
Zhang and Mueller (2005) present sensitivity (A) and bias (b) metrics that would be suitable for use with this measure. These metrics are based on Signal Detection Theory (SDT) but correct for limitations of SDT assumptions by allowing for hit rates and false alarm rates to be 1 or 0, as well as correct a formula error in an earlier metric. Sensitivity (A) is a non-parametric measure of the difference between responses to signal (correct elements) and response to noise (incorrect elements). Bias (b) is a tendency to respond one way or another (e.g., to select all elements or not select any).

Identifying Relationships

Awareness of Elements is an analytical ability (breaking the problem down into smaller parts), but synthesis (putting the elements together to form a whole) is also required to recognize the entire problem space. In the Identifying Relationships module, the test taker begins synthesis by mapping the relationships between elements and identifying the exchange of inputs and outputs with the system's environment. This module requires the test taker to generate causal statements to describe dyadic relationships between important elements in the system (see Figure 7).

Figure 7

Interface for Identifying Relationships



To form the causal statements, the test taker is presented with a list of elements and a list of verb phrases and asked to create statements structured as follows: [Element 1] [Relational verb / verb phrase] [Element 2]. The list for each element consists of the full list of elements relevant to the system. Once an element is selected for Element 1 or Element 2, it cannot be selected for the other element in the same sentence (i.e., no self-reference). The test taker will also select a verb that best fits the relationship between the two selected elements (e.g., increases, decreases, feeds into, is fed by, etc., as appropriate). Selecting two elements and a causal verb between the two elements creates a causal sentence. The assessment captures how many causal sentences are made, as well as the number of correct and incorrect causal statements made. Correct sentences will move scores in the positive direction and incorrect sentences will move scores in the negative direction. The instructions will specify that only direct relationships will be counted as correct. Relationship statements that are indirect but correct have been identified and will not increase or decrease scores on the module. The number of possible relationship statements and correct relationship statements for each subsystem are shown in Table 4. The test taker must complete the Identifying Relationships module and submit their answers before moving to the next module.

Table 4

Number of Identifying Relationships Items for Each Subsystem

	Elements	Verbs	Correct Relationships	Possible Relationships	% Correct of all Possible
Elevator Call Room Access	10	4	9	96	9.4%
Elevator Auto-Mode	8	2	8	112	7.1%
Sentinels Modes and Uses	14	1	9	182	4.9%
Sentinel Mode Change	4	2	8	24	33.3%
Reactor Activation	7	3	6	126	4.8%
Reactor Containment	5	3	6	60	10.0%
Swarmbot Zones	6	5	11	180	6.1%
Swarmbot Rules of Individual Behavior	4	3	7	36	19.4%
Swarmbot Rules of Collective Behavior	8	2	11	112	9.8%

Understanding System Dynamics

After the test taker has finalized their responses to Identifying Relationships, they must complete the Understanding System Dynamics module by answering questions about feedback loops and basic system archetypes. Earlier in the assessment when the test taker first receives instructions on the systems to explore through the Mission Console, the test taker is asked to "calibrate" software that will ostensibly analyze the information provided about the system and send that information to Earth. While presumably conducting the calibration task, the test taker will actually be completing the first portion of the Understanding System Dynamics module. This enables the software to first provide the test taker with the basic language needed for describing feedback loops as either balancing (negative) or reinforcing (positive), as well as providing a basic familiarity with system archetypes (modified feedback loops). Following this introduction to system dynamics, the test taker is shown simple feedback loops and system archetypes and asked to categorize them. The test taker answers questions of increasing complexity about system archetypes and complex system dynamics. There are currently 10 multiple choice questions in this portion of the test.

Later, when test takers are completing the modules for each test system, the module poses additional Understanding System Dynamics questions for each system. These questions either require the user to select the option that best describes the types of feedback loops involved in a specific system or subsystem (labeled Understanding Feedback Loops) or to choose the systems diagram that best depicts that system's dynamics (labeled Understanding Causal Loop Diagrams). Table 5 provides a breakdown of the number of items of both types for each test system. All the Understanding System Dynamics items across systems and including the calibration activity are included in the score. All items are multiple choice and scored as correct or incorrect. Total scores are the percent correct across all Understanding System Dynamics items. Once the Understanding System Dynamics module is completed for all systems, the test takers are told that their insights will be conveyed back to Earth leaders to prepare additional strategies.

Table 5

Number of Understanding System Dynamics Items for Each System

	Understanding Feedback Loops Items	Understanding Causal Loop Diagram Items
Elevators	2	1
Sentinels	3	1
Reactors	4	2
Swarmbots	3	1
Total	12	5

Evaluate and Revise

A concept for an Evaluate and Revise module is planned for future versions of the systems thinking assessment. Once the test taker has submitted answers to the Understanding System Dynamics modules for all of the systems, the test taker would be told that there is additional information on some of the systems. The test taker would return to a designated site for each of the systems and look for more information. For time efficiency, a button will be provided that takes the test taker to a specific location of the ship to look for the additional information. For some systems, information that was not available before is made available; other systems have no new information. The Evaluate and Revise dimension would be measured through the test taker making appropriate updates to the Awareness of Elements and Identifying Relationships modules based on the new information. The test taker would collect points for identifying new system elements and identifying correct relationships. They would lose points for adding elements that are not part of the system or identifying incorrect relationships.

There are also other possibilities for measuring this dimension, such as collecting within-simulation measures. Evaluating and revising one's understanding of the system is a continuous process, making continuous data collection an appealing measurement approach. Future data collections will explore the feasibility of using behavioral indicators for Evaluate and Revise.

Application to Mission

In the Application to Mission module, the test taker is prompted to answer questions about the relevance of the system, its structure, its similarity to other systems, and ways in which that information might be useful. In this module, the test taker must leverage their exploration and model building activities and use reasoning about their understanding of the systems to provide a suitable course of action for furthering the mission (rescuing prisoners, escaping the ship, helping Earth forces defeat the alien threat). Scores on this module will be assessed using multiple choice situational judgment questions about each of the systems that are explored in the Mission Console. The number of questions for each system can be seen in Table 6. Each question contains four to six response options and asks test takers to select the best one. If test takers have a better understanding of the systems, they should be better judges of the best courses of action. An overall Application to Mission score is calculated from the average across systems.

Table 6

Number of Application to Mission Items for Each System

	Number of Prompts	Number of Rated Statements
Elevators	2	10
Sentinels	2	11
Reactors	2	12
Swarmbots	2	13
Total	8	46

Revisions to Initial Antecedent Measures

Adis et al. (2017) developed measures for five constructs (hierarchical working memory capacity, pattern recognition, spatial ability, cognitive flexibility, and cognitive complexity), and assessed construct validity using participants from Amazon's Mechanical Turk platform (sample sizes ranging from 119 to 137). Table 7 describes each of the measures. Based on the results from Adis et al. (2017) and the input of project team SMEs, modifications and corresponding programming changes were made, as needed. Items that were not successful in the validation were dropped, revised, or replaced, and parameters of measures that had limited variability in scores were modified. See Adis et al. (2017) for descriptions of initial score distributions and available psychometrics.

Hierarchical Working Memory Capacity

Hierarchical working memory capacity was developed in the form of a traditional location-number memory binding task with the novel addition of a three-level hierarchy for the location component (Adis et al., 2017). The distribution closely followed a normal distribution aside from the lower tail being truncated by the lower bound of the test, which suggested scoring was too difficult at the lower levels of the test. To address the difficulty, the project team changed the measure to start at a memory span of one instead of two and changed the scoring to give credit for partial recall (i.e., remembering only the correct location or number) rather than only giving credit if both location and number are recalled.

Table 7

Description of Five Previously Developed Antecedent Constructs Developed

Construct	Measurement
Hierarchical Working Memory Capacity	Multi-level information processing test – complex memory span task where items are at different levels of a hierarchy
Spatial Ability	a) Dynamic shape searching b) 3-D shape slicing c) Tilted container fluid levels d) Picture perspective-taking
Cognitive Flexibility	a) Flexibility: Scenario-based hypothesis formation–change when new information is provided b) Micro/Macro Switching – reaction times for switching from micro to macro levels on the Pattern Recognition task
Pattern Recognition	a) Anomaly detection approach where subjects identify anomalies in a factory-like environment b) Multi-level symbol recognition using Navon figures as stimuli
Cognitive Complexity	Dynamic sorting task – drag and drop items into groups and label them

Spatial Ability

Four measures of spatial ability were developed by Adis et al. (2017): (a) intrinsic-static spatial ability was assessed with an object search task consisting of four items that each had a unique shape to be found up to eight times within 45 seconds, (b) intrinsic-dynamic spatial ability was assessed with 25 multiple choice mental cutting items, (c) extrinsic-static spatial ability assessed participants' judgment of the fullness of a jar tilted at different angles and (d) extrinsic-dynamic spatial ability was assessed with 12 items in which the participant had to identify the location of a photographer and the direction of the camera. Adis et al. found that the distribution for intrinsic-static spatial ability was truncated on the tail end with higher scores. This suggested that the items were not difficult enough to effectively assess the highest performing test-takers. Thus, the project team added two more difficult items to the measure. The distributions of both the intrinsic-dynamic and extrinsic-dynamic measures were truncated at the front and tail end, which suggested that scoring was too difficult to effectively assess the lowest performing test-takers at the lower levels of the test, and not difficult enough to effectively assess the highest performing test-takers at the higher levels of the test. Thus, more difficult and easier items were developed for each measure to improve the distributions. Moreover, some items that did not perform well were removed from each measure (e.g., items with low item-total correlation). No changes were made to the extrinsic-static spatial ability test from the earlier version.

Cognitive Flexibility

There are two measures of cognitive flexibility (see Adis et al., 2017): (a) scenario-based hypothesis formation was assessed with 20 cognitive flexibility scenarios (CF1), and (b) micro-macro task switching was assessed using reaction times for switching from micro to macro levels on the pattern recognition task (CF2). The format of the CF1 measure was changed to drag-and-drop with numerous response options and an area for write-in guesses. This will facilitate future automated scoring. Five of the most promising scenarios were selected to remain in the CF1 measure, based on the number, distribution, and flexibility of responses collected from CF1. The CF2 measure was shortened from 13 blocks with 10 trials each to 13 blocks with 8 trials each, resulting in a 20 percent reduction in test length.

Pattern Recognition

Two pattern recognition instruments were developed by Adis et al. (2017): (a) anomaly detection using an assembly line minigame (PR1) and (b) multilevel pattern recognition using Navon figures as stimuli (PR2). The anomaly detection pattern recognition test yields two scores: a count of the number of errors detected (corrected for false alarms as described above) and the response latency until errors are found, for all found errors. The distribution of errors caught was very positively skewed (.73) as was the distribution of latencies (1.88). The positive skew indicated that a large number of participants did not catch any or only a few errors. To address the non-normal distributions, we slowed down the rate of the assembly line, which should make it easier to detect errors. No changes were made to the multilevel pattern recognition items; however, graphic updates were made to improve item display.

Cognitive Complexity

The cognitive complexity measure consisted of three different sorting activities: (a) a set of objects sorted based on how they look (CC1), (b) the same set of objects sorted based on how they are used (CC2), and (c) a set of systems-related words or phrases sorted by appearance, function, or any other conceivable way (CC3). No changes were made to this measure from the earlier version.

Integration Into the Systems Thinking Assessment

These existing antecedent assessments were incorporated into the storyline under the guise of mental testing being conducted on the humans by the aliens. The test taker must keep up the appearance of being under mind control in order to blend in; therefore, they must complete the assessments in the mind control tests to the best of their ability. The first five times that the test taker is captured by patrolling Sentinels, they will be taken to a testing center on the ship and directed to the updated antecedent measures. After data have been collected on each measure, the captured test taker will simply resume play shortly after being captured.

Development of Additional Antecedent Measures

Measures were developed for the three additional systems thinking related antecedent attributes (curiosity, creativity, and openness to information) through test taker actions, reactions, and responses to questions during the systems thinking assessment. Table 8 provides a summary of the attribute definitions and measurement approaches.

Table 8

Description of Three Additional Systems Thinking Antecedents

Construct	Definition	Measurement
Creativity	Producing ideas or products that are novel or unusual and provide a useful approach or perspective in a given situation.	Interactive scenario responses – novelty, fluency, and usefulness of ideas generated to solve a problem
Curiosity	Possessing the desire to know or learn about people, things, and relationships and taking action to investigate them.	Interactive scenario responses – measuring amount of time spent and extent of investigating unusual objects
Openness to Information	Taking note of new information or points of view that are relevant to models or plans and considering ways in which the information impacts those models or plans.	Responses to questions that test participants' perceptiveness for and awareness of issues, characteristics, or features of the ship that are not explicitly addressed but are observable for those who are looking.

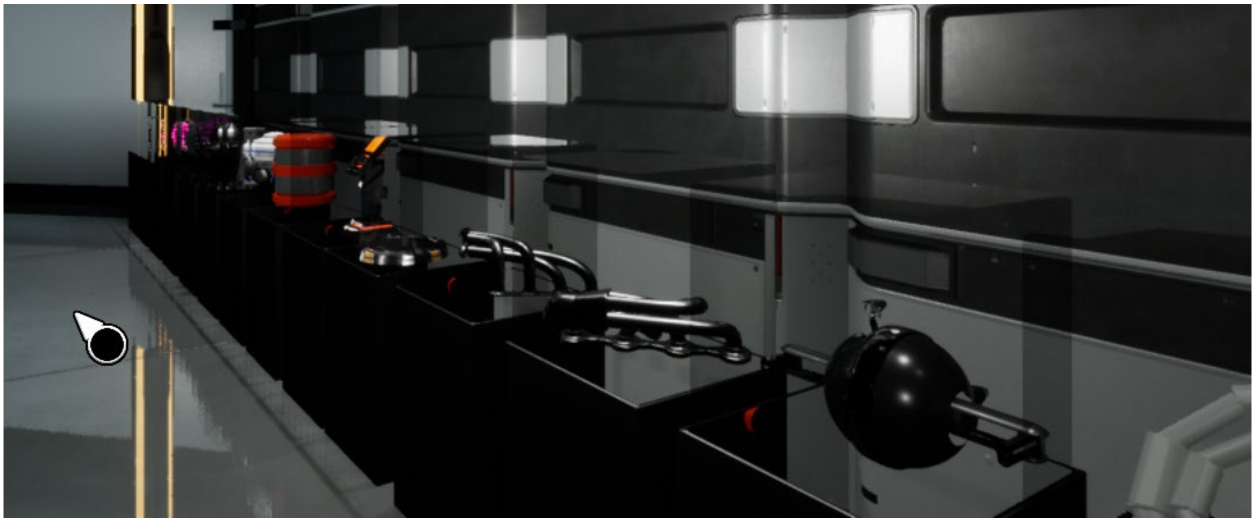
Curiosity

Curiosity is measured using interactive scenarios in which the test taker is asked to wait in a room that contains 26 alien artifacts and has no mission-related activities to accomplish during that period. Over 30 pre-existing graphic objects were acquired that were: (a) available to include, and (b) ambiguous enough to pass as some type of alien object. Researchers fabricated a name and 4-5 pieces of information or characteristics for each object.

The test taker is taken to this waiting room three times during the session, for a two-minute period each time (see Figure 8 for a screenshot of the waiting room). During that time period, the test taker has the option to (a) investigate the objects, (b) play a distractor task (a puzzle game that is available in the room), or (c) take no action during that period. The premise is that test takers who are more curious will spend more time and effort to examine the objects available in the room, rather than doing nothing or playing the distractor game. By making the test taker wait in the room for a specific period of time, we ensure that they do not have to choose between curiosity behaviors and behaviors associated with investigating the systems, which could reflect decision-making or time management rather than just curiosity.

Figure 8

Screenshot of Curiosity Objects in Waiting Room



The test taker will be given the instructions: “Wait here for your next test. Your test will start in 2 minutes. You can play this game if you want while you wait.” The test taker can then navigate to the puzzle game or the museum area. The alien items in the museum are in cases where the test taker can click on a button to learn more information about each artifact. Each time they click they learn additional information, with a maximum of 4-5 pieces of information available per object. Objects cannot be removed from their cases.

The raw data collected from the assessment include the amount of time spent on each object, the number of clicks for additional information on each object, and the number and identity of objects visited. The curiosity assessment provides indices such as the number of times a test taker returns to an object after the first view, the total time spent on objects, the average depth of information viewed within an object, and the overall breadth of information viewed across all objects. Many of these are expected to be highly correlated. We will examine each indicator alone and in combination with other indicators in order to determine the most efficient and effective measurement approach.

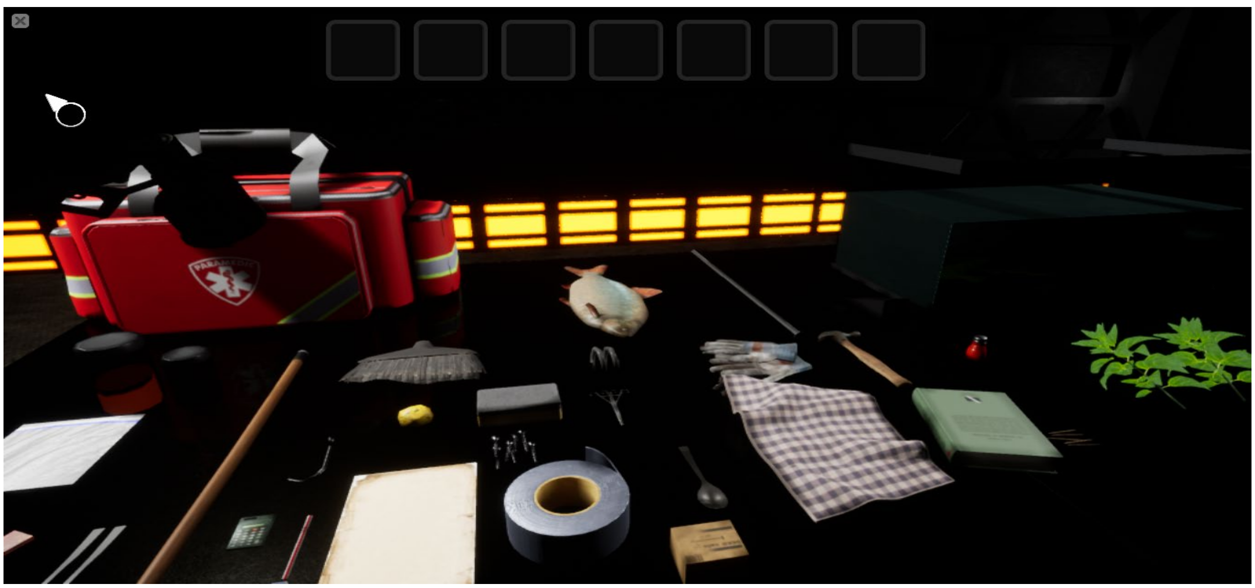
Creativity

To measure creativity, the project team developed concepts for several interactive scenarios that present test takers with situations in which they need to generate multiple ways to accomplish a task. The test takers are given two minutes to generate as many possible creative responses as they can. The responses generated by the participant are then scored on several criteria: the number of ideas generated, the uniqueness of the ideas, and their feasibility. Details about the object properties are openly available to the test taker so that they are being tested on their creativity in applying the objects rather than their curiosity or openness to learning about the objects.

The project team developed initial concepts for three scenarios. For example, in one scenario, the test taker is taken back to a jail cell after they are recaptured. The test taker notices some objects (e.g., a paperclip, pen) placed around the cell (see Figure 9 for a screenshot of objects from one scenario). The test taker must use the objects, either on their own or in combination with other objects, to form a strategy to escape from the cell. Once an object or combination of objects is selected, the test taker must specify where and how that object would be used in the present situation. They are allowed to sequence together multiple combinations of objects, locations, and uses to form their overall solution. When test takers are finished crafting their solution to the problem, they submit it and can begin to create another solution. There is a range of possible combinations for the objects and a range of possible applications of the object combinations on the different locations in the scenario (e.g., various potential escape routes in the jail cell scenario).

Figure 9

Screenshot of Creativity Scenario Objects



There are three possible scoring dimensions for creativity: fluency, originality, and feasibility. During the initial testing, the test taker's score will be based only on their fluency, which is captured through the number of solutions they generate. Future iterations of the creativity measure will incorporate scoring for the originality and feasibility of solutions as well. Solutions will be evaluated on originality by comparing each solution to other solutions that have been generated throughout the scenario's history and scaling originality scores to be inversely proportional to the relative frequency of previously generated solutions. Solutions that have never occurred before would earn maximum originality points and solutions that have occurred with high frequency would earn few originality points. In order to derive feasibility scores, each solution component (object-location-use) will be rated by project personnel on its feasibility from 0 to 1. Once a sufficient library of solutions has been generated and rated on feasibility, the feasibility assessment can be automatically scored. Overall solutions will be given feasibility scores which are the product of the feasibility of each component in the solution.

Because of the experimental nature of the measure, both in stimulus delivery and scaling approach, the present version of the systems thinking assessment includes only one creativity scenario. Using this one scenario, the project team will evaluate the viability of the approach for measuring creativity prior to developing the other two scenarios. Once all three scenarios are incorporated in the assessment, a method will be developed to derive an overall creativity index using all three scoring dimensions and all three scenarios. An example of an index that incorporates all three scoring dimensions would be a point system in which test takers receive three points for each solution rated as "high" in originality, two points for solutions rated as "moderate" in originality, and one point for solutions rated as "low" in originality. Scores could then be adjusted for feasibility, subtracting two points for solutions rated as "very low in feasibility" and one point for solutions rated as "low in feasibility." Solutions rated as moderate or high in feasibility will not be adjusted.

Openness to Information

Openness to information is measured using questions that test participant assumptions about bigger contextual issues. Test takers are instructed to gather information that will be helpful to understand how the U.S. forces can either successfully attack the ship or enable the prisoners to escape. These instructions suggest that they should pay attention to information that relates to these issues. Throughout the course of the session, observations, objects, and information are available in the environment of the ship. Near the end of the test, the test taker is presented with several questions regarding characteristics of the ship. Test takers that are higher in openness to information are expected to have noticed this information and thus will be more likely to answer more questions correctly, earning a higher score on this measure.

A total of 42 questions were initially developed. These were evaluated for the extent to which the information needed to answer the question is observable in the developed software. The questions that pertained to information that was readily observable in the current version of the software were selected for inclusion, resulting in a 23-item measure. Each item is scored by comparing test takers' ratings with an average of SME ratings. Total scores will be composed of overall averages across all items.

An alternative measurement approach for openness to information was initially developed in which test takers' information capturing, collecting, and exploring behaviors resulted in the accumulation of openness to information points. While the question-based measurement approach was selected for inclusion, the point-based approach might be able to supplement the question-based measure. Numerous covert data sources are available from test takers' behaviors during the assessment. These could possibly be used in a dynamic way to determine whether a given question should be included in the openness to information score; that is, if a test taker did not view observations relevant to a specific question while exploring the ship, that question would not be used in determining the openness to information score. By collecting metrics on the locations visited and the views that were rendered, we can quantify how many opportunities for openness to information observations were made. This would allow us to create a separate index of openness to information that scales responses to items based on the number of observations that were made. The feasibility of alternative measurement approaches for openness to information will be explored after preliminary data are collected. Using supplementary data from covert measures, the research team will be able to compute additional openness to information indices and examine distributions and correlations to other metrics.

Next Steps

With a simulation-based assessment, there is always more that can be done to optimize measurement, functionality, or the graphic experience of the test taker. Work will continue on these user experience aspects of the systems thinking assessment. Additionally, initial validation research is needed to examine psychometric properties of the systems thinking dimensions and antecedents. These data will be essential for assessing the reliability of the measures as well as the descriptive statistics and distributions of the scores the measures produce. Analyses will also be conducted to identify ways to shorten the assessment, including eliminating measures that are redundant or non-performing, or eliminating items to shorten the length of individual measures. These data will provide insight regarding adjustments that can be made to improve the effectiveness of each test system and the assessment and scoring approaches. Each of the four systems will be evaluated to determine whether test takers were able to make the appropriate observations and complete the dimension assessments. Test taker data will be critical for evaluating the effectiveness of the creativity measure and determining if additional creativity scenarios should be developed. Measures such as the Application to Mission dimension and openness to information will benefit from insights obtained from user data to determine where additional test taker observations are needed.

Correlations between the planned measures and behavioral indicators of activities will be examined. These indicators may be useful to leverage as stealth assessment approaches that tacitly capture individual differences based on the exploration or interaction patterns of test takers. For example, a stealth measure of Awareness of Elements may be obtainable from data such as the time taken to explore particular elements or areas. Once sufficient data have been collected, exploratory investigations can be conducted into the identification of behavioral indicators that reflect the test takers' awareness of system elements. An approach was mentioned previously for the openness to information measure using stealthily collected indicators to modify openness to information based on opportunities of each test taker, as another example. If reliable and valid estimates can be obtained from stealth measures, future versions of the systems thinking assessment may be able to omit the explicit response formats, like the element selection interface for Awareness of Elements in the Mission Console or the direct questions for openness to information. The assessment platform already collects large amounts of data and if many of the constructs lend themselves to stealth data collection, this could reduce testing time by using only stealth indicators. Sorting through these data and correlating behavioral indicators with measurement data will be time-consuming but may result in new measurement techniques and/or a shorter assessment battery.

Conclusions

This research centered on the development of an innovative tool to assess individual differences that are conceptually related to systems thinking. Concepts were developed for simulation-based assessments that measure eight antecedent constructs and five dimensions potentially related to systems thinking performance. A fictional context was designed that integrates the assessments within a systems thinking simulation.

In addition, the assessment can measure multiple behavioral indicators of the systems thinking dimensions derived from interactions between the test taker and the systems within the environment. By collecting data and identifying behavioral indicators, it may be possible to shorten the test by making it less driven by responses to questions and more driven by exploratory behavior. Though this research is ongoing, it introduces an exciting and promising approach and a measurement approach for systems thinking that is novel and highly needed in the field of systems thinking.

Moreover, this research informs the practices of identifying, selecting, and/or classifying personnel within the Army. Numerous jobs in the Army require Soldiers to work with or within systems. For example, Wheeler (1999) argued for the importance of systems thinking down to the level of team leaders and encouraged an understanding of teams as systems. Other literature emphasizes the importance of systems thinking for strategic leaders (e.g., Allen et al, 2009; Macaulay, 2011) or views systems thinking is a component of strategic thinking (e.g., Sackett et al., 2016). Given the role of systems thinking across Army echelons, tools for quantifying systems thinking potential could be extremely valuable.

References

Adis, C., Wisecarver, M., Key-Roberts, M., Hope, T., Pritchett, C. S., & Henderson, J. (2013). Building sociocultural systems thinking capabilities. (Project Report, W5J9CQ-13-C-0001). U.S. Army Research Institute for the Behavioral and Social Sciences.

Adis, C., Wisecarver, M., Raber, C., Wind, A. P., & Canali, K. G. (2017). *Innovative Tools to Assess Systems Thinking Ability* (Technical Report 1362). U.S. Army Research Institute for the Behavioral and Social Sciences. https://apps.dtic.mil/sti/pdfs/AD1045468.pdf

Allen, C. D., Cunningham, G. K., & Klinger, J. (2009). *Systems Thinking for Strategic Leaders*. Army War College.

Braun, W. (2002). The system archetypes. *System*, 1-27.

Cabrera, D. (2006). *Systems thinking* [Unpublished doctoral dissertation]. Cornell University. https://ecommons.cornell.edu/bitstream/handle/1813/2860/DerekCabreraDissertation.pdf;sequence=1

Cabrera, D., Colosi, L., & Lobdell, C. (2008). Systems thinking. *Evaluation and Program Planning, 31*(3), 299-310.

Canali, K., Wind, A., & Willford, J. (2017). *Cyber Selection Test Research Effort for U.S. Army New Accessions*. International Military Testing Association. https://apps.dtic.mil/dtic/tr/fulltext/u2/1044605.pdf

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge University Press.

Checkland, P. (1981). *Systems thinking, systems practice*. John Wiley & Sons Inc. https://doi.org/10.1002/1099-1743(200011)17:1+<::AID-SRES384>3.0.CO;2-N

Conway, A. R. A., Kane, M. J., & Engle, R. W. (2003). Working memory capacity and its relation to general intelligence. *Trends in Cognitive Sciences, 7*, 547-552.

Dedmond, W. K. (2014). *Systems Thinking: A Way to Understand Socio-political Environments*. Army Command and General Staff College. https://apps.dtic.mil/dtic/tr/fulltext/u2/a614191.pdf

Dennis, M. B. (2010). *Systems thinking and design: Making learning organizations a reality in the US Army*. Army Command and General Staff College. https://pdfs.semanticscholar.org/2bc4/d33598d187f12f305b877dfdd701047dd0aa.pdf

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. A. (1999). Working memory, short-term memory, and general fluid intelligence: A latent variable approach. *Journal of Experimental Psychology: General, 128*(3), 309-331.

Flood, R. L., & Jackson, M. C. (1991). Creative problem solving: Total systems intervention. In M. C. Jackson (Ed.), *Systems Methodology for the Management Sciences* (pp 271–276). Springer. https://doi.org/10.1007/978-1-4899-2632-6_11

Guilford, J. P. (1967). *The nature of human intelligence*. McGraw-Hill.

Hambrick, D. Z., & Engle, R. W. (2003). The role of working memory in problem solving. In J. E. Davidson & R. J. Sternberg (Eds.), *The psychology of problem solving* (pp. 176–206). Cambridge University Press. https://doi.org/10.1017/CBO9780511615771.007

Jackson, M. C., & Keys, P. (1984). Towards a system of systems methodologies. *Journal of the Operational Research Society*, *35*(6), 473–486. https://doi.org/10.1057/jors.1984.101

Kim, K. H. (2005). Can only intelligent people be creative? A meta-analysis. *Journal of Secondary Gifted Education, 16,* 57-66.

Macaulay, D. (2011). *Systems Thinking: A Foundational Art for Strategic Leaders*. Army War College.

Meadows, D. H. (2008). *Thinking in systems: A primer* (D. Wright, Ed.). Chelsea Green Publishing.

Midgley, G. (1997). Mixing methods: Developing systemic interventions. In J. C. Mingers & A. Gill (Eds.), *Multimethodology: The theory and practice of combining management science methodologies*. Wiley.

Midgley, G. (Ed.). (2003). *Systems thinking*, (Vol. 1-4). Sage Publications.

Mingers, J. C. (2006). *Realising systems thinking: Knowledge and action in management science*. Springer. https://doi.org/10.1108/03684920810885017

Pretz, J. E., Naples, A. J., & Sternberg, R. J. (2003). Recognizing, defining, and representing problems. In J. E. Davidson & R. J. Sternberg (Eds.), *The psychology of problem solving* (pp. 3-10). Cambridge University Press. https://doi.org/10.1017/CBO9780511615771.002

Raine, A., Reynolds, C., Venables, P. H., & Mednick, S. A. (2002). Stimulation seeking and intelligence: A prospective longitudinal study. *Journal of Personality and Social Psychology, 82,* 663-674.

Randle, J. M., & Stroink, M. L. (2012). *Systems thinking cognitive paradigm and its relationship to academic attitudes and achievement*. Unpublished Honors Thesis. Lakehead University. https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.864.4778&rep=rep1&type=pdf

Rimfield, K., Shakeshaft, N. G., Malanchini, M., Rodic, M., Selzam, S., Schofield, K., Dale, P. S., Kovas, Y., & Plomin, R. (2017). Phenotypic and genetic evidence for a unifactorial structure of spatial abilities. *PNAS, 114(10),* 2777-2782.

Sackett, A. L., Karrasch, A. I., Weyhrauch, W. S., & Goldman, E. F. (2016). *Enhancing the strategic capability of the Army: An investigation of strategic thinking tasks, skills, and development.* (Research Report 1995). U.S. Army Research Institute for the Behavioral and Social Sciences.

Senge, P. M. (1990). *The fifth discipline: The art and practice of the learning organization*. Random House. https://doi.org/10.1002/pfi.4170300510

Shute, V. J. (2011). Stealth assessment in computer-based games to support learning. *Computer Games and Instruction*, *55*(2), 503–524. http://myweb.fsu.edu/vshute/pdf/shute%20pres_h.pdf

Shute, V. J., Masduki, I., & Donmez, O. (2010). Conceptual framework for modeling, assessing and supporting competencies within game environments. *Technology, Instruction, Cognition & Learning*, *8*(2), 137-161.

Shute, V. J., Wang, L., Greiff, S., Zhao, W., & Moore, G. (2016). Measuring problem solving skills via stealth assessment in an engaging video game. *Computers in Human Behavior*, *63*, 106–117. https://doi.org/10/f82qxd

Silva, P. J. (2015). Intelligence and creativity are pretty similar after all. *Educational Psychology Review, 27,* 599-606.

Sweeney, L. B., & Sterman, J. D. (2000). Bathtub dynamics: Initial results of a systems thinking inventory. *System Dynamics Review: The Journal of the System Dynamics Society*, *16*(4), 249–286. https://doi.org/10.1002/sdr.198

Sweeney, L. B., & Sterman, J. D. (2007). Thinking about systems: Student and teacher conceptions of natural and social systems. *System Dynamics Review: The Journal of the System Dynamics Society*, *23*(2–3), 285–311. https://doi.org/10.1002/sdr.366

Thibodeau, P. H., Frantz, C. M., & Stroink, M. L. (2016). Situating a measure of systems thinking in a landscape of psychological constructs. *Systems Research and Behavioral Science*, *33*(6), 753–769. https://doi.org/10.1002/sres.2388

Von Stumm, S., Hell, B., & Chamorro-Premuzic, T. (2011). The hungry mind: Intellectual curiosity is the third pillar of academic performance. *Perspectives on Psychological Science, 6,* 574-588.

Wallach, M. A., & Kogan, N. (1965). *Modes of thinking in young children: A study of the creativity–intelligence distinction*. Holt, Rinehart, & Winston.

Wheeler, T. A. (1999). *U.S. Army Rapid Deployment Task Force Dynamics-Rapid Teaming.* Army Command and General Staff College, School of Advanced Military Studies.

Williams, J. D. (2007). *Is the U.S. Army a Learning Organization?* Army War College. https://apps.dtic.mil/dtic/tr/fulltext/u2/a469631.pdf

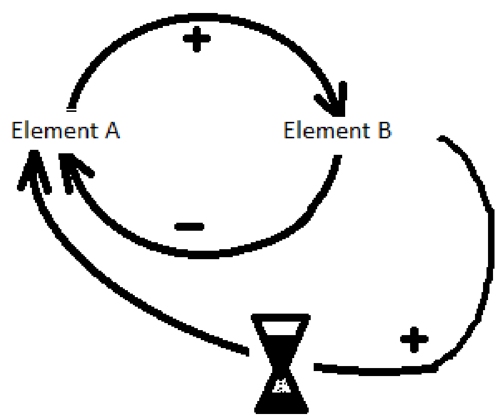
Zhang, J., & Mueller, S. T. (2005). A note on ROC analysis and non-parametric estimate of sensitivity. *Psychometrika, 70*(1), 203–212. doi:10.1007/s11336-003-1119-8

APPENDIX A. SYSTEM CONCEPT TEMPLATE

Overarching System: System X	
System Archetype: Fixes that Fail	**Key Subsystems:**
Brief Description of System:	**Key Mechanisms:**
Functionality:	
System Missions (prompts to direct test takers to investigate system):	
System Interactions and Observations	
A.	Trigger: Observations: Added Elements: Actions:
B.	Trigger: Observations: Added Elements: Actions:
C.	Trigger: Observations: Added Elements: Actions:
Awareness of Elements (AE)	
Embodying System: • Element A* • Element B+ • Element C (s) * indicates key element in subsystem + indicates hidden element when looking at system (s) indicates a situational factor or other non-physical element	

How the test taker can learn about each system element?	
Element A	Location; observation
Element B	Location; observation
Element C	Location; observation
Identifying Relationships	
Key Relationships: **Embodying System:**	
Summary of Terms and Verbs:	
How the test taker can learn about each system relationship?	
Action	Description
Understanding System Dynamics	
Archetype Type and Description: Fixes that fail: Fixes that fail is a system archetype that in system dynamics is used to describe and analyze a situation, where a fix effective in the short-term creates side effects for the long-term behavior of the system and may result in the need of even more fixes.	

Diagram:



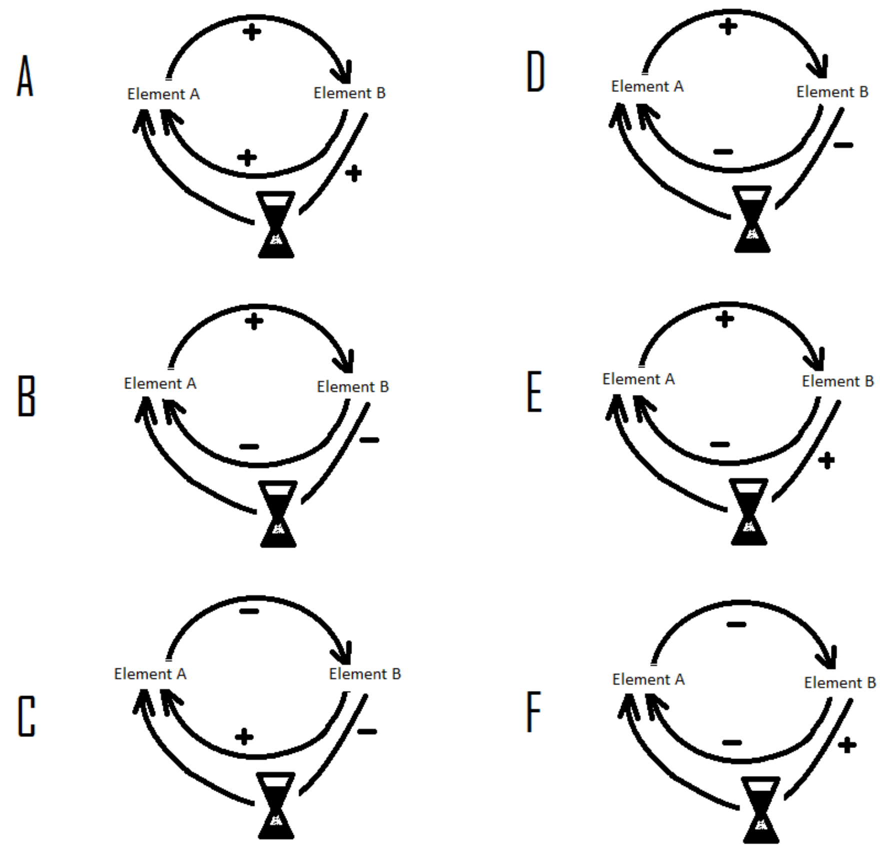
Understanding System Dynamics Relationships NOT in Identifying Relationships:	
USD Prompts	1. You may have noticed that XXX has a limitation. When you do this action, a resulting action occurs. Given this, the overall situation is: a. A reinforcing loop b. A balancing loop c. A combination of loops*

d. Not a feedback loop

2. If there was no XXX, the situation would be best characterized as:

   a. A reinforcing loop*
   b. A balancing loop
   c. A combination of loops
   d. Not a feedback loop

Which of the following figures most accurately shows the relationships involved when …. ?



Evaluate & Revise

Now to update our understanding of the system:

After further consideration, I have determined that there is more to learn about the system, and I believe key information can be found in …

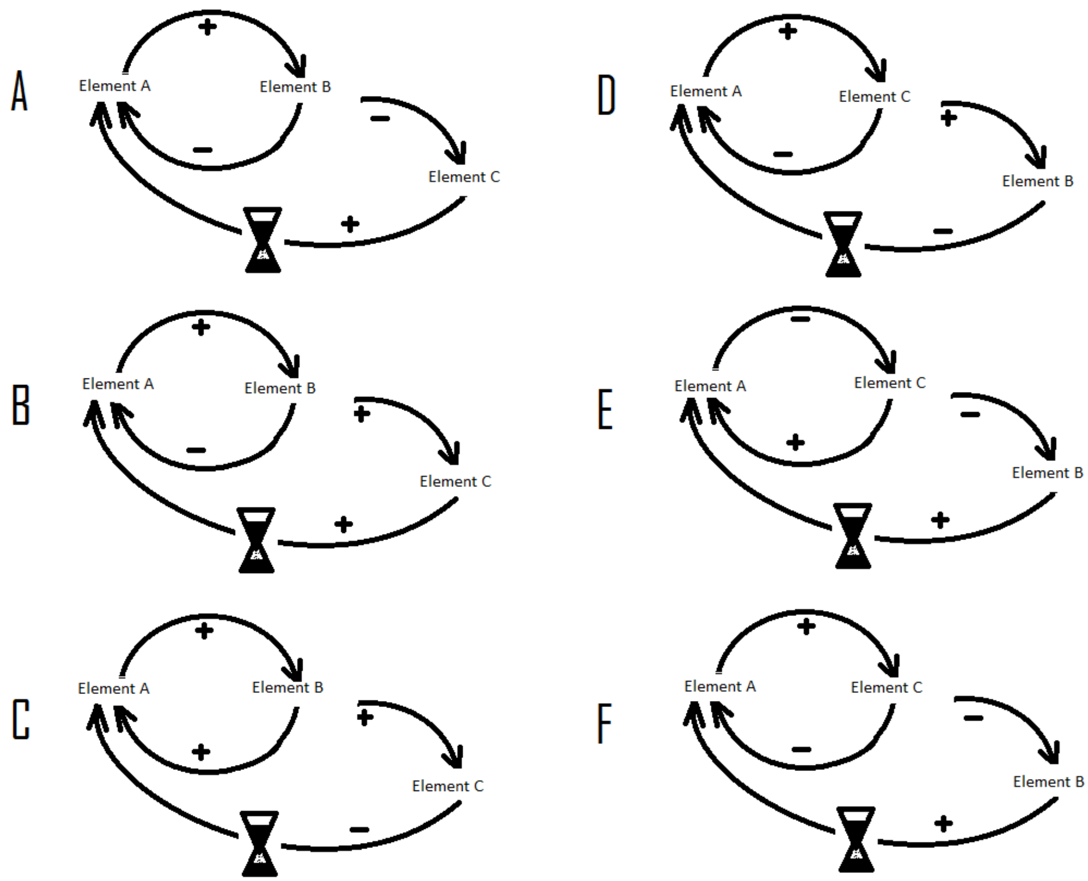
Hidden Aspects for Evaluate/Revise:

Added elements for *Awareness of Elements* revision:

Added relationships for *Identifying Relationships* revision:

***Understanding System Dynamics* revisions:**

Based on what you have learned, which of the figures most accurately shows the relationships involved when …



Application to the Mission
SJT Questions: