AFRL-RH-WP-TR-2019-0043

USE OF TECHNOLOGY-ENHANCED WORK SIMULATIONS FOR CYBER ASSESSMENT

Michael D. Coovert
Jaclyn Martin
David J. Howard
Frances Kim
Rachel Dreibelbis
Matthew S. Arbogast
Sean Potter

MDC & Associates

January 2019
Interim Report

DISTRIBUTION STATEMENT A: Approved for public release.

AIR FORCE RESEARCH LABORATORY
711TH HUMAN PERFORMANCE WING,
AIRMAN SYSTEMS DIRECTORATE,
WRIGHT-PATTERSON AIR FORCE BASE, OH 45433
AIR FORCE MATERIEL COMMAND
UNITED STATES AIR FORCE

NOTICE AND SIGNATURE PAGE

Using Government drawings, specifications, or other data included in this document for any purpose other than Government procurement does not in any way obligate the U.S. Government. The fact that the Government formulated or supplied the drawings, specifications, or other data does not license the holder or any other person or corporation; or convey any rights or permission to manufacture, use, or sell any patented invention that may relate to them.

This report was cleared for public release by the 88th Air Base Wing Public Affairs Office and is available to the general public, including foreign nationals. Copies may be obtained from the Defense Technical Information Center (DTIC) (http://www.dtic.mil).

AFRL-RH-WP-TR-2019-0043 HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION IN ACCORDANCE WITH ASSIGNED DISTRIBUTION STATEMENT.

//signature//

THOMAS R. CARRETTA
Work Unit Manager
Supervisory Control and Cognition Branch

//signature//

JOSEPH P. NALEPKA
Chief, Supervisory Control and Cognition Branch
Warfighter Interface Division

//signature//

LOUISE A. CARTER
Chief, Warfighter Interface Division
Airman Systems Directorate

This report is published in the interest of scientific and technical information exchange, and its publication does not constitute the Government's approval or disapproval of its ideas or findings.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

The public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Department of Defense, Washington Headquarters Services, Directorate for Information Operations and Reports (0704-0188), 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to any penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number. **PLEASE DO NOT RETURN YOUR FORM TO THE ABOVE ADDRESS**.

1. REPORT DATE *(DD-MM-YY)*
21-01-19

2. REPORT TYPE
Interim

3. DATES COVERED *(From - To)*
1/18/2018-2/28/2019

4. TITLE AND SUBTITLE
Use of Technology-Enhanced Work Simulations for Cyber Assessment

5a. CONTRACT NUMBER
FA8650-14-D-6500-0007

5b. GRANT NUMBER
Not applicable

5c. PROGRAM ELEMENT NUMBER
52202F

6. AUTHOR(S)
Michael D. Coovert, Jaclyn Martin, David J. Howard, Frances Kim, Rachel Dreibelbis, Matthew S. Arbogast, Sean Potter

5d. PROJECT NUMBER
5329

5e. TASK NUMBER
09

5f. WORK UNIT NUMBER
H0SA (532909TC)

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
MDC & Associates
4004 Stanley Rd.
Plant City, FL 33565

8. PERFORMING ORGANIZATION REPORT NUMBER
MDC18-19-AF1

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Air Force Materiel Command
Air Force Research Laboratory
711th Human Performance Wing
Airman Systems Directorate
Warfighter Interface Division
Supervisory Control & Cognition Branch
Wright-Patterson AFB, OH 45433

10. SPONSORING/MONITORING AGENCY ACRONYM(S)
711 HPW/RHCI

11. SPONSORING/MONITORING AGENCY REPORT NUMBER(S)
AFRL-RH-WP-TR-2019-0043x

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release. 88ABW-2019-0557. Cleared 28 August 2019

13. SUPPLEMENTARY NOTES
Subcontract number: FPH02-S025/180123; Report contains color.

14. ABSTRACT
There were three objectives to this effort: 1) identify aptitudes and traits required for success in select enlisted and officer Air Force cyber careers using archival information, 2) identify which cyber aptitudes and traits can be measured through existing DoD tests, 3) provide a summary of relevant literature and recommendations for how serious games could be used to measure cyber aptitudes and traits, including those where gaps in assessment currently exist. We identified an extensive list of aptitudes and traits and established the extent to which each is adequately assessed. Five traits (Analytical Thinking, Adaptability, Dependability, Persistence, Situational Awareness) and four aptitudes (Active Learning, Decision Making, Deductive Reasoning, Systems Thinking) are the person characteristics most related to on the job cyber performance for which there is currently inadequate assessment and could be assessed via technologically-enhanced work simulations such as with a serious game.

15. SUBJECT TERMS
Competency, mental competency, trait, aptitude, personnel selection, personnel development, serious game, cybersecurity, cyber science, computing related activities, technology

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT:	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON (Monitor)
Unclassified	Unclassified	Unclassified	SAR	136	Thomas R. Carretta **19b. TELEPHONE NUMBER** *(Include Area Code)* (937) 713-7143

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std. Z39-18

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

PREFACE

This report was prepared under Subcontract to Infoscitex (IST) (Subcontract NO. HIRT FPH02-S022, Prime Contract N0. FA8650-14-D-6500).

We thank Mark Rose, Tom Carretta, and Annette Rizer for their input and assistance throughout the project. We thank Infocitex for the administrative support. We thank the numerous subject matter experts that took time to speak with us and to complete and distribute the survey used for this research. We appreciate their time and thoughtful responses. Finally, we thank Nicole Stoneley, Stephanie Boettcher, and Joelle Castle for their assistance with the report.

SUMMARY

The United States Air Force (USAF) performs extensive research across several domains to support and enhance the warfighter. One mission of the Human Effectiveness Directorate in the Air Force Research Laboratory (AFRL) is to provide the Air Force human-centered research in order to "…optimize and protect the airman's capabilities to fly, fight, and win in air, space, and cyberspace" (USAF, 2016). Working in support of this directive for the Air Force Personnel Center/Strategic Research and Assessment group, our focus is on the use of technology-enhanced work simulations for cyber assessments.

There were three objectives to this effort. The first objective was to identify aptitudes and traits required for success in select enlisted and officer Air Force cyber careers using archival information. The second objective was to identify which cyber aptitudes and traits can be measured through existing Department of Defense (DoD) tests. The third objective was to provide a summary of relevant literature and recommendations for how serious games could be used to measure cyber aptitudes and traits, including those where gaps in assessment currently exist.

Objective 1 summary: Much archival information covering cyber occupations was provided by the Air Force and reviewed by our team. The information proved valuable in establishing the knowledge, skills, abilities, and traits important to both enlisted and officers in cyber occupations. We extended the scope of the Air Force cyber occupations through our examination of cyber careers in the Occupational Information Network (O*NET). Since O*NET's coverage is based on occupations in the U.S. economy, it brings substantial breadth to the coverage of knowledge, skills, abilities, and other characteristics (KSAO) requirements for cyber jobs. We see from this source that six cognitive ability factors subsuming nineteen specific abilities are deemed important for cognitive workers. Furthermore, seven different work style factors covering sixteen individual facets are also important for cyber occupations. The work by University of Maryland Center for Advanced Study of Language (UMD CASL) is also insightful for abilities to consider for the cyber occupations. Gap analysis provides a mapping of aptitudes and traits to specific Air Force Specialty Codes (AFSCs). Interviews with subject matter experts (SMEs) were most helpful in that they both confirmed required aptitudes and traits we had uncovered from the archival analyses, and provided insights into other aptitudes and traits to consider. After the list of important aptitudes and traits was finalized, SME input into the relative importance of each construct was integral to determining the critical constructs for each cyber position of interest.

Objective 2 summary: We reviewed the potential best bets for measuring the cyber aptitudes and traits that were identified as important in Objective 1. We first examined the eight existing DoD tests identified by the USAF as potentially useful assessments for assessing cyber aptitudes and traits. Information from the AF provided reports and materials was summarized to provide an overview of the relevant content and psychometrics for each of the DoD tests. The cyber aptitudes and traits identified in Objective 1 were then rationally sorted into DoD tests that measured similar aptitudes and traits. This sorting was used to identify gaps of aptitudes and traits where there were no current DoD measures that measured corresponding aptitudes/traits (see Table 30 for the list of gaps in current assessment). The second section in this document

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 28 August 2019

reviews potential alternative measures for each of the gaps listed in Table 30. These include tests from other federal government agencies, tests from commercial vendors, and concepts based on the scientific literature. We determined from this review that five traits (Analytical Thinking, Adaptability, Dependability, Persistence, Situational Awareness) and four aptitudes (Active Learning, Decision Making, Deductive Reasoning, Systems Thinking) are the person characteristics most related to on the job cyber performance for which there is currently inadequate assessment.

Objective 3 summary: Technologically-enhanced assessment for cyber operators could occur through the use of a serious game. We review serious games and features common across games and identify existing classifications of serious games. Issues of performance assessment during game play are discussed and due to pragmatic considerations, it would be most beneficial to have a game that does not rely on evaluators monitoring performance during play. We describe popular game engines that would be suitable for the current task and describe four example scenarios that could provide the context for the assessment of the traits and aptitudes identified in Objective 2. The process of statistical assessment of the traits and aptitudes should be carried out via confirmatory factor analysis, and multi-trait-multi-method modeling. Finally, usability analysis of the game needs to establish it is sound from a human factors and human-computer interaction perspective.

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 28 August 2019

1.0 INTRODUCTION

Recent advances in technology has led to a surge in interest surrounding the use of serious games for providing immersive and realistic simulations that may allow for the elicitation and ultimate measurement of constructs that are difficult to measure through traditional assessments (Fetzer, 2015; Landers, Auer, Collmus, & Armstrong, 2018; Tippins, 2015). Serious games are characterized by features, such as narrative guides that further task completion, feedback and rewards, and by technology-enhanced user interface and experience (Tippins, 2015). Consideration of alternative methods of assessment for complex constructs is especially important as the nature of work continues to change, requiring increasingly complex KSAOs (Brannick, Pearlman, & Sanchez, 2017).

One instance where the changing nature of work pushes change in selection systems is for the selection of cyber-security workers. With a projected 3.5 million unfilled cyber-security jobs in 2021 (NeSmith, 2018) combined with a shortage of qualified workers in the market, there is increasing interest in moving past traditional assessments to identify those with the aptitude to succeed in the cyber realm. Moreover, with the demand for cyber workers outpacing the supply of cyber workers, employers need to increase the applicant pool through moving focus from technical knowledge and certifications to more distal predictors of job performance, such as abilities and traits.

The limited scientific literature on abilities and traits important for cyber-security job performance is conceptual and theoretical in nature with no predictive validation data currently available. The non-technical abilities and traits proposed to be important for cyber-security job performance are systems thinking ability, analytical thinking, active or continuous learning, communication, integrity/civic duty, time management/selective attention/vigilance, and attention to detail (Dawson & Thomson, 2018; Dreibelbis, Martin, Coovert, & Dorsey, 2018; Jose, LaPort, & Trippe, 2016).

Some of these constructs are difficult to measure with traditional self-report methods. Aside from allowing for the measurement of more complex constructs, serious games increase psychological fidelity, which allows for greater face validity of the assessment and can increase candidates' perceptions of fairness in the selection process. Serious games could also increase engagement for applicants, which could result in less applicant attrition and more positive candidate reactions when compared to traditional assessment. Moreover, serious games could allow for the measurement of multiple constructs simultaneously, which may ultimately reduce testing time.

The present work aims to identify aptitudes and traits that are important for cyber-security positions, to identify existing measures that are available for the important aptitudes and traits, and to evaluate the utility of serious games for measuring the important aptitudes and traits in a selection context.

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 28 August 2019

2.0 OBJECTIVE 1 – IDENTIFY APTITUDES AND TRAITS

Our work to identify the aptitudes and traits required for competency in various cyber AFSCs began with a review of the military research on testing and assessment. Next. we considered scientific outlets to garner a complete picture of the required individual attributes. A gap analysis was done in order to focus future work on those aptitude and trait measures that require development. As a validation check, these aptitudes and traits were reviewed by subject matter experts to determine their relative importance to the Air Force cyber career fields.

2.1 Review of Air Force Archival Materials

Our literature review included select published articles, papers, technical reports, military publications and doctrine, as well as briefings made available to us by the USAF and other military sources. First, we reviewed the provided materials, paying careful attention to relevant aptitudes and traits needed for success in the six career fields of interest (enlisted: 3D0X2 - Cyber Systems Operations; 3D0X3 - Cyber Surety; 1B4X1 - Cyber Warfare Operations; 1N4X1A - Digital Network Analyst; officer: 17DX - Cyberspace Operations; 17SX - Cyber Warfare Operations).

The Air Force Enlisted Classification Directory (AFECD; Air Force Personnel Center, 2017a) contained brief job descriptions and entry requirements for the enlisted career fields. Similarly, the Air Force Officer Classification Directory (AFOCD; Air Force Personnel Center, 2017b) provided brief job descriptions and entry requirements for the officer cyber career fields. Combined, the AFECD and AFOCD documents listed the duties, responsibilities, and specialty qualifications relevant to all levels of each cyber career field (Superintendent, Journeyman, Apprentice, Helper). For the 3D0X2 and 3D0X3 positions, Occupational Analysis Report (OAR) information was also available for use to identify knowledge, skills and abilities (KSAs) important for the career field (Lambert, 2014a, 2014b). Summary information from these materials are detailed in Table 1, sorted by AFSC.

Notably, for the officer career fields, the duties and responsibilities, specialty qualifications, and related specialties are nearly identical. Specifically, there are no differences listed in the AFOCD between the specialty qualifications and entry requirements between the 17DX Network Operations 17SX Cyber Warfare Officers. However, the AFOCD lists the following duties and responsibilities under the 17DX Cyberspace Operations officer career field (but not under 17SX Cyber Warfare Officers):

1. Reviews terrain and weather information
2. Translates operational requirements into architectural and technical solutions. Works with commanders to deliver complete capabilities that include technical and procedural components. Researches or oversees research of technologies and advises commanders on associated risks and mitigation factors in conjunction with meeting requirements.
3. Directs extension, employment, reconfiguration, adaptation and creation of portions of cyberspace to assure mission success for combatant commanders. This includes both deliberate and crisis action scenarios.

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 28 August 2019

Table 1. KSA Information Obtained from Air Force Archival Materials

AFSC	KSA Information	Source
Cyber Systems Operations (3D0X2)	• Knowledge is mandatory of cyber systems elements; capabilities, functions, and technical methods for system operations.	AFECD
	• Minimum Mechanical, Administrative, General, Electrical (MAGE) composite or Alternate Minimum MAGE with Cyber Test Scores: General Score (Arithmetic Reasoning [AR] + Verbal Expression [VE]): 64; or General Score (AR + VE): 54 and Cyber-Test 60	AFECD
Cyber Surety (3D0X3)	• Knowledge is mandatory of information systems (IS) resources; capabilities, functions and technical methods for IS operations; organization and functions of networked IS resources; communications-computer flows, operations and logic of electromechanical and electronics IS and their components, techniques for solving IS operations problems; and IS resources security procedures and programs including Internet Protocols.	AFECD
	• Minimum MAGE or Alternate Minimum MAGE with Cyber Test Scores: General Score (AR + VE): 64; or General Score (AR + VE): 54 and Cyber-Test 60	AFECD
Cyber Warfare Operations (1B4X1)	• Knowledge is mandatory of computer operating systems, software applications, and hardware components. In addition, understanding of networking fundamentals to include protocols, network addressing, and, network infrastructure to include telecommunications theory and data communications. Must be proficient on wireless networking and understand cryptography to include utilization and exploitation techniques. Must have understanding of applicable laws governing cyber operations.	AFECD
	• Minimum MAGE Score: General Score (AR + VE): 64	AFECD
Digital Network Analyst (1N4X1A)	• Must gain and maintain knowledge of global communications procedures; analytical techniques; organization of the national intelligence structure; intelligence organizations and systems; Information Operations; organization of designated military forces; geography; collection and reporting, systems, principles, methods, and procedures; effective writing principles; oral and written intelligence information presentation; and directives for handling, disseminating, and safeguarding classified information.	AFECD
	• Minimum MAGE or Alternate Minimum MAGE with Cyber Test Scores: General Score (AR + VE): 62; or General Score (AR + VE): 57 and Cyber-Test 60	AFECD
Cyber-space Operations (17DX)	• Knowledge is mandatory including electronics theory, information technology, telecommunications and supervisory and control systems including cryptography, vulnerability assessment and exploitation techniques. Additionally, knowledge will include operational planning, governing cyberspace operations directives, procedures and tactics.	AFOCD
	• For entry into the career field, it is mandatory that an officer accession meet at least one of the tiers listed in the Career	AFOCD

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 28 August 2019

	Intermission Program (CIP) matrix for the career field in which they are to be accessed. Each career field has a different set of requirements and tier structure. If an officer does not meet any of the tier requirements in the matrix, they are not qualified to enter the career field. Tier 1 (mandatory): Computer and Information Sciences and Support Services or Computer Engineering or Engineering Physics/Applied Physics or Industrial Engineering or Electromechanical Engineering or Electrical, Electronics and Communications or Electrical, Electronics and Communications Engineering Technology/Technician or Computer Technology/Computer Systems Technology or Cyber/Electronics Operations and Warfare or Mathematics and Computer Science or Accounting and Computer Science or Computational Science or Management Information Systems or Mathematics. Tier 2 (desired): Engineering or Engineering Technologies or Mathematics and Statistics or Physics or Chemistry. Tier 3 (permitted): Any Degree.	
Cyber Warfare Operations (17SX)	• Knowledge is mandatory including electronics theory, information technology, telecommunications and supervisory and control systems including cryptography, vulnerability assessment and exploitation techniques. Additionally, knowledge will include operational planning, governing cyberspace operations directives, procedures and tactics.	AFOCD
	• For entry into the career field, it is mandatory that an officer accession meet at least one of the tiers listed in the CIP matrix for the career field in which they are to be accessed. Each career field has a different set of requirements and tier structure. If an officer doesn't meet any of the tier requirements in the matrix, they are not qualified to enter the career field. Tier 1 (mandatory): Computer and Information Sciences and Support Services or Computer Engineering or Engineering Physics/Applied Physics or Industrial Engineering or Electromechanical Engineering or Electrical, Electronics and Communications or Electrical, Electronics and Communications Engineering Technology/Technician or Computer Technology/Computer Systems Technology or Cyber/Electronics Operations and Warfare or Mathematics and Computer Science or Accounting and Computer Science or Computational Science or Management Information Systems or Mathematics. Tier 2 (desired): Engineering or Engineering Technologies or Mathematics and Statistics or Physics or Chemistry. Tier 3 (permitted): Any Degree.	AFOCD

As can be seen in Table 1, comprehensive descriptions of the cyber occupations of focus are maintained by the Air Force. We next examined cyber occupations found in the non-military sector of the US economy by investigating databases maintained by the US government.

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 28 August 2019

2.2 Cyber Careers in Civilian Occupations and Other Agencies

Developed under the sponsorship of the US Department of Labor/Employment and Training Administration (USDOL/ETA), the Occupational Information Network (O*NET; https://www.onetonline.org/) contains information on hundreds of occupations and describes working in the United States today. O*NET provides essential information for each occupation, including coverage for the skills and knowledge required of job incumbents; as well as the abilities, interests, and work styles found in those same individuals.

One very useful feature of O*NET is called Crosswalk. The feature allows veterans to identify occupations that involve similar tasks and KSAOs as those used in their military occupation. Employing Crosswalk, we identified 20 civilian occupations that are similar to the AFSCs that are the focus of our work. Utilizing R software, we performed web scraping of those 20 occupations in order to compile a complete list of tasks and KSAOs in civilian occupations that are similar to the six Air Force occupations that are the focus of our work. This involved setting up script in R to search the crosswalk for each of the six AFSCs ("3D0X2", "3D0X3", "1B4X1", "1N4X1A", "17DX", "17Sx") and requesting the relevant O*NET data from each corresponding civilian job.

To identify important aptitudes for the cyber AFSCs, we included both the abilities and work styles identified as essential for the related civilian occupations (see Tables 2 and 3). Research supports that “abilities are relatively enduring basic capacities for performing a wide range of different tasks” (Carroll, 1993; Fleishman & Reilly, 1992; N. G. Peterson et al., 2001, p. 457). Although extremely useful, abilities do not tell the whole story on worker characteristics or requirements. As such, we also examined work styles. O*NET defines work styles as “personal characteristics that can affect how well someone performs a job.” We included the O*NET work styles to evaluate those as possible traits important for cyber positions. We also included the 14 aptitudes and traits found in the UMD CASL studies previously determined to be important for cyber AFSCs (see Table 4; O’Rourke, Karuzis, S. Kim, et al., 2017).

We first present findings from O*NET for abilities followed by the findings for work styles. Subsequently, the aptitudes and traits from CASL are presented.

2.2.1. Abilities listed in O*NET for cyber occupations

The abilities in Table 2 are arranged according to the number of occupations out of 20 in O*NET which reported those as important to the occupation. O*NET reports the importance weights as a number ranging from 0-100. We computed the mean importance rating for each ability by averaging across the cyber occupations in which each is found. A criticality score is then computed by multiplying the averaged importance rating by the number of occupations in which it is required (these are reported in Table 2 and Figure 1). Figure 1 depicts the abilities as arranged in the abilities taxonomy employed by O*NET and defined in Fleishman and Reilly (1992).

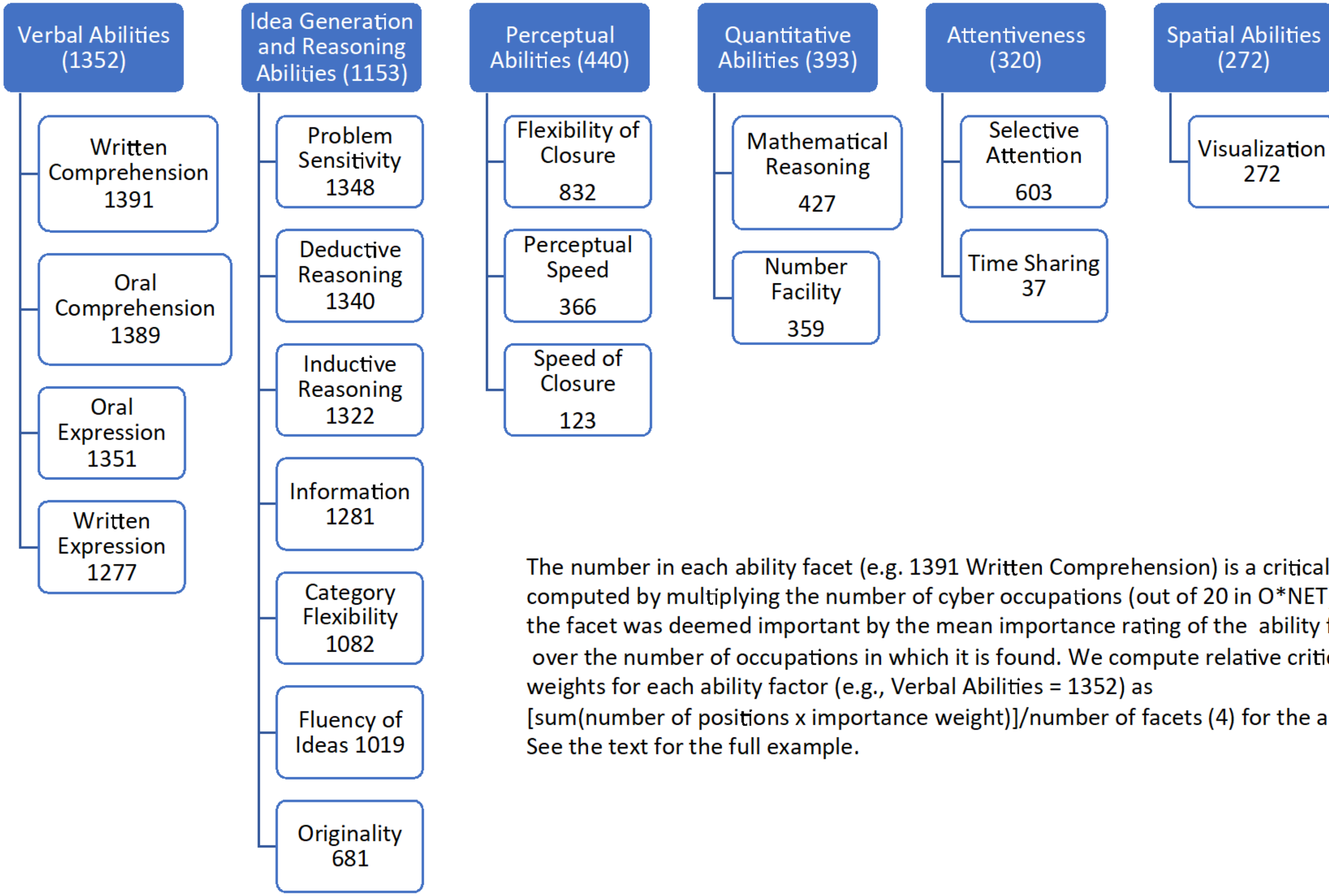
Table 2. O*NET Abilities from 20 Similar Civilian Cyber Positions

Ability	Definition (The ability to…)	# POS	IMP	Critic
Written Comprehension	Read and understand information and ideas presented in writing.	19	73.24	73.24
Oral Comprehension	Listen to and understand information and ideas presented through spoken words and sentences.	19	73.12	1389
Oral Expression	Communicate information and ideas in speaking so others will understand.	19	71.12	1351
Problem Sensitivity	Tell when something is wrong or is likely to go wrong. It does not involve solving the problem, only recognizing there is a problem.	19	70.94	1348
Deductive Reasoning	Apply general rules to specific problems to produce answers that make sense.	19	70.53	1340
Inductive Reasoning	Combine pieces of information to form general rules or conclusions (includes finding a relationship among seemingly unrelated events).	19	69.59	1322
Information Ordering	Arrange things or actions in a certain order or pattern according to a specific rule or set of rules (e.g., patterns of numbers, letters, words, pictures, mathematical operations).	19	67.41	1281
Written Expression	Communicate information and ideas in writing so others will understand.	19	67.24	1277
Category Flexibility	Generate or use different sets of rules for combining or grouping things in different ways.	19	56.94	1082
Fluency of Ideas	Come up with a number of ideas about a topic (the number of ideas is important, not their quality, correctness, or creativity).	18	56.59	1019
Flexibility of Closure	Identify or detect a known pattern (a figure, object, word, or sound) that is hidden in other distracting material.	16	52.00	832
Originality	Come up with unusual or clever ideas about a given topic or situation, or to develop creative ways to solve a problem.	13	52.41	681
Selective Attention	Concentrate on a task over a period of time without being distracted.	12	50.24	603
Mathematical Reasoning	Choose the right mathematical methods or formulas to solve a problem.	9	47.41	427
Perceptual Speed	Quickly and accurately compare similarities and differences among sets of letters, numbers, objects, pictures, or patterns. The things to be compared may be presented at the same time or one after the other. This ability also includes comparing a presented object with a remembered object.	8	45.76	366
Number Facility	Add, subtract, multiply, or divide quickly and correctly.	8	44.82	359
Visualization	Imagine how something will look after it is moved around or when its parts are moved or rearranged.	6	45.35	272

Speed of Closure	Quickly make sense of, combine, and organize information into meaningful patterns.	3	40.88	123
Time Sharing	Shift back and forth between two or more activities or sources of information (such as speech, sounds, touch, or other sources).	1	37.29	37

*Note. # POS = Number of positions (out of 20). IMP = Mean importance rating (0-100). Criticality = #POS * IMP.*

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 28 August 2019



The number in each ability facet (e.g. 1391 Written Comprehension) is a criticality score computed by multiplying the number of cyber occupations (out of 20 in O*NET) in which the facet was deemed important by the mean importance rating of the ability facet averaged over the number of occupations in which it is found. We compute relative criticality weights for each ability factor (e.g., Verbal Abilities = 1352) as [sum(number of positions x importance weight)]/number of facets (4) for the ability factor. See the text for the full example.

Figure 1. Cognitive abilities for cyber occupations described in O*NET

As summarized in Figure 1, six different cognitive ability factors (categories) capture the 19 distinct cognitive abilities found in civilian cyber occupations. These factors are verbal, idea generation and reasoning, perceptual, attentiveness, quantitative, and spatial. To help determine the relative importance of these aptitude factors we computed a relative criticality score as follows. Each ability factor has one or more facets (e.g., verbal has 4: written comprehension, oral comprehension, oral expression, and written expression). Each facet has two numbers associated with it, one indicating the number of cyber occupations in O*NET where the facet is found; and the second indicating the mean importance of that facet to the occupations (these are reported in Table 2 and Figure 1). We computed a mean criticality weight for each type of ability by multiplying the number of occupations times the importance weight (from O*NET). This provided a relative weighting as a function of the number of cyber occupations in which it is found and the importance weight for that ability as computed in O*NET. We computed this individually for each of the 19 abilities. A factor level criticality was also computed by summing those criticality values within ability factor, and then dividing by the number of facets for the ability factor. For example, Verbal Abilities = [written comprehension (19*73.24)+oral comprehension (19*73.12)+oral expression (19*71.12)+written expression (19*67.24)]/4 which results in 1,352.

This number allows us to form a numerical assessment of the importance of each ability factor to the cyber occupations. We then standardized these by computing z-scores to help with interpretation. Using this relative comparison, two ability factors [verbal (z=1.48); idea generation and reasoning (z=1.06)] were much more important than the other four [perceptual (z=-.46), quantitative (z=-.56), attentiveness (z=-.71), and spatial (z=-.81)]. Figure 1 depicts this relative importance. While this is useful, it is important to point out two things regarding this relative comparison. First, all six ability factors are important in cyber occupations with the numerical comparison allowing an evaluation of relative criticality at the factor level for the abilities. Second, this is merely one possible way to compute relative criticality, but it appears useful for our purposes.

2.2.2. Work styles listed in O*NET for cyber occupations

We now move to the second major source of information from O*NET and that is work styles. As with abilities, we performed web scraping for the work styles associated with the 20 civilian cyber occupations identified via the O*NET crosswalk feature. Table 3 reports the work styles, number of occupations each style is descriptive of, the mean importance rating (range is 0-100), and a criticality score.

Table 3. Work Styles Culled from O*NET for Civilian Cyber Occupations

Work Style	Definition (Job Requires…)	# POS	IMP	CRIT
Attention to Detail	Being careful about detail and thorough in completing work tasks.	17	89.82	1525
Dependability	Being reliable, responsible, and dependable, and fulfilling obligations.	18	86.53	1555
Integrity	Being honest and ethical.	18	86.41	1549
Analytical Thinking	Analyzing information and using logic to address work-related issues and problems.	18	85.24	1539

Initiative	A willingness to take on responsibilities and challenges.	17	80.41	1363
Cooperation	Being pleasant with others on the job and displaying a good-natured, cooperative attitude.	16	80.24	1281
Adaptability/Flexibility	Being open to change (positive or negative) and to considerable variety in the workplace.	17	78.94	1341
Persistence	Persistence in the face of obstacles.	15	78.06	1173
Stress Tolerance	Accepting criticism and dealing calmly and effectively with high stress situations.	17	76.24	1293
Achievement/ Effort	Establishing and maintaining personally challenging achievement goals and exerting effort toward mastering tasks.	15	75.24	1131
Self-Control	Maintaining composure, keeping emotions in check, controlling anger, and avoiding aggressive behavior, even in very difficult situations.	15	73.94	1105
Independence	Developing one's own ways of doing things, guiding oneself with little or no supervision, and depending on oneself to get things done.	15	72.24	1086
Innovation	Creativity and alternative thinking to develop new ideas for and answers to work-related problems.	15	68.24	1024
Leadership	Willingness to lead, take charge, and offer opinions and direction.	15	67.24	998
Concern for Others	Being sensitive to others' needs and feelings and being understanding and helpful on the job.	15	62.94	939
Social Orientation	Preferring to work with others rather than alone, and being personally connected with others on the job.	15	56.00	830

*Note. # POS = Number of positions (out of 20). IMP = Mean importance rating (0-100). CRIT = Criticality= #POS * IMP.*

Examination of Table 3 and Figure 2 reveal that 16 work styles are critical to cyber occupations, and these can be organized under seven specific work style factors or categories. Following the procedure described for aptitudes, we computed a relative criticality score for each work style. See Figure 2. Again, standardizing the criticality score to z scores for further interpretation we found the following. Based on the relative criticality, the most influential work style categories are conscientiousness ($z = 1.82$), followed by practical intelligence (.44), adjustment (.25), and achievement orientation (.12). Furthermore, we found independence (-.60), interpersonal orientation (-.96), and social influence (-1.07) to be somewhat less critical.

O*NET has proved to be a rich source of information for cyber occupations. One particular strength is that it cuts across all cyber occupations in the US economy, those in both the public and private sectors. In all, 19 separate cognitive abilities are identified as being critical (seen

Figure 1). These are organized under six cognitive ability factors, with the relative importance of the factors also presented in Table 2 and Figure 1. Similarly, 16 specific work styles were deemed important to incumbents in cyber occupations. As with abilities, the criticality of individual work styles was also computed and can be found in Table 3 and Figure 2, along with the relative importance of work styles at the factor level.

We now turn to a third source of information, that provided by the University of Maryland CASL work which is feeding into a new test for cyber aptitude and talent assessment (CATA).

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 28 August 2019

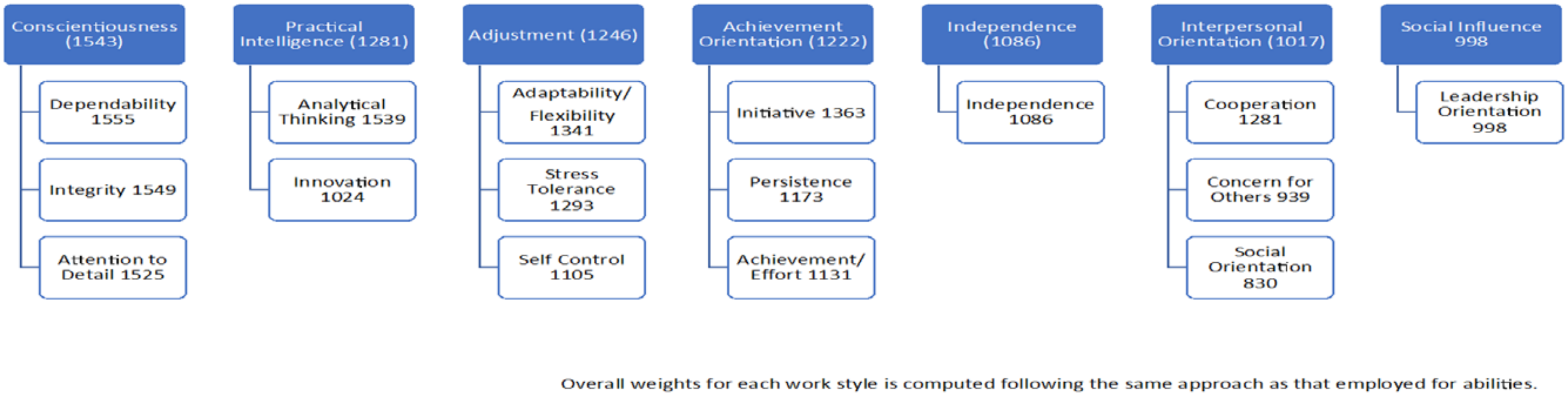


Figure 2. Work styles for cyber occupations described in O*NET

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 28 August 2019

2.2.3. Aptitudes and traits as measured by University of Maryland CASL for cyber occupations

As mentioned above, in addition to examining civilian occupations for aptitudes and traits important to cyber positions, we also looked at the work done by those developing a new test for cyber assessment for the Air Force (O'Rourke, Karuzis, S. Kim, et al., 2017). These aptitudes and traits are listed in Table 4.

Table 4. CASL Aptitudes and Traits Important for Cyber Positions

Aptitude/Trait	Definition
Anomaly Detection	The ability to detect information that is anomalous in a larger context, such that it does not conform to the expected pattern.
Pattern Recognition and Scanning	The ability to scan incoming information, detect patterns and react quickly.
Complex Problem Solving	The ability to learn and effectively manipulate systems, which are complex, opaque, and dynamic (Frensch & Funke, 2014).
Rule Induction	The ability to determine the rules that govern a pattern.
Need for Cognitive Closure*	The need to arrive at a solution during problem solving
Need for Cognition*	The degree to which individuals enjoy participating in mentally demanding tasks.
Tolerance for Risk*	The likelihood of an individual to be risk taking or risk averse, and is a factor known to influence decision making (e.g., Bechara, Damasio, & Damasio, 2000).
Spatial Visualization	The ability to form and manipulate visuospatial representations.
Visuospatial Working Memory	The workspace for briefly holding and manipulating information from the spatial domain (Baddeley, 1986, 2007; Baddeley & Hitch, 1974).
Psychomotor Speed	The ability to respond quickly and to control the speeded motor response in the face of interference.
Resistance to Interference	The ability to respond quickly and accurately in the face of proactive interference (Monsell, 1978; Sternberg, 1966).
Convergent Creative Thinking	The ability to explore a variety of solutions by forming connections between concepts that are typically weakly related or unrelated, and ultimately to hone in on one single, correct solution to a problem (see Cropley, 2006, for a review).
Mental Model Ability	The ability to construct abstract, internal representations of a situation, real or imagined, derived from a narrative or other form of input (Ehrlich & Johnson-Laird, 1982) and provide a basis for inference making and successful recall of information (Zwaan & Radvansky, 1998).
Vigilance	The ability to remain vigilant or sustain attention during a task that occurs over a prolonged period.

Note. Table is adapted from information from O'Rourke et al. (2017). The constructs with an asterisk (*) are treated as traits, while all others are treated as abilities.

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19

The UMD CASL work provides a set of aptitudes and traits important for assessing cyber personnel. As can be seen from the table, many of these are unique from those of O*NET.

Prior to leaving this section, we note that our team did reach out to other governmental agencies (intelligence and other) in an attempt to garner information regarding the selection approach they employ for cyber personnel. These agencies were largely non-responsive to our requests. We were, however, able to obtain some insights regarding their procedures from comments made by SMEs we interviewed. These were captured in the SME section below.

2.3 Gap analysis: Comparison of the AFSC and Civilian Information

Since abilities and traits are less situation-specific and variable over time compared to knowledge and skills (DuVernet, Dierdorff, & Wilson, 2015), they are the most useful for the current objective. Given that the archival material included information on KSAOs for the different AFSCs, we needed to translate that material into abilities and traits (like those detailed above from O*NET and CASL) in order to determine what is currently covered in the positions. From the details in the archival materials, we used rational sorts to match the archival information on tasks and KSAs into the aptitudes and traits listed in Tables 2, 3, and 4. These occupational details were drawn primarily from the AFECD and AFOCD documents, which contained the majority of the information regarding important aptitudes and traits. These documents were supplemented by the CFETPs, and for only the 3D0X2 and 3D0X3 positions, we also reviewed information from the OARs.

Some KSAs that did not fit into any of the specified aptitudes and traits, thus we reviewed the literature for applicable constructs. Specifically, knowledge of information systems or information technology or telecommunications was mentioned for each of the AFSCs, so Information and Technology Aptitude (Trippe, Russell, Schwartz, & Weissmuller, 2008) was identified as a potentially important aptitude.

Of the 31 aptitudes identified as important for cyber positions, 23 were covered in some capacity in the archival Air Force occupation materials. Eight aptitudes were explicitly recommended in the archival materials for at least half of the jobs. These aptitudes included Information and Technology Aptitude, Anomaly Detection, Complex Problem-Solving, Visualization, Vigilance, and Pattern Recognition. Of the 23 covered aptitudes, sixteen were either explicitly documented for less than three positions or were implied by the documentation to be an important aptitude for more than one position. In Table 5, implied aptitudes are presented with asterisks.

The remaining eight abilities identified in civilian positions were not accounted for by the Air Force positions (about 26% of the examined abilities). These eight abilities not identified in the archival material were Category Flexibility, Convergent Creative Thinking, Mental Model Ability, Originality, Psychomotor Speed, Perceptual Speed, Speed of Closure, and Resistance to Interference. There were several abilities that had poor coverage in the occupational materials, namely Deductive Reasoning, Flexibility of Closure, Information Ordering, Mathematical Reasoning, Number Facility, Oral Comprehension, Oral Expression, Selective Attention, and Time Sharing. Oral Expression, Oral Comprehension, Deductive Reasoning, Information Ordering, and Category Flexibility are especially noteworthy as the

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19

Table 5. Summary of Important Aptitudes Covered in Archival Materials.

Aptitude	3D0X2	3D0X3	1B4X1	1N4X1A	17DX	17SX
[1]Anomaly detection	*	X	X	*	X	X
[2]Category Flexibility						
[1]Complex problem-solving	X	X	X	*	X	X
[1]Convergent Creative Thinking						
[2]Deductive Reasoning	*					
[2]Flexibility of closure	*		*	*	*	
[2]Fluency of Ideas	*		*		X	X
[2]Inductive Reasoning/ [1]Rule Induction	*	X				
Information and Technology Aptitude	X	X	X	X	X	X
[2]Information Ordering		*		*		
[2]Mathematical Reasoning	*			*	*	*
[1]Mental Model Ability						
[1]Modeling program execution				*	X	X
[2]Number Facility	*			*	*	*
[2]Oral Comprehension	*	*	*	*	*	*
[2]Oral Expression	*	*	*	*	*	*
[2]Originality						
[1]Pattern recognition and scanning	X			X		X
[2]Perceptual Speed						
[2]Problem Sensitivity		X		*	*	*
[1]Psychomotor Speed						
[1]Resistance to Interference						
[2]Selective Attention	*	*	*	*	*	*
[1]Spatial Visualization/[2]Visualization	X	X	X	*	X	X
[2]Speed of Closure						
[2]Time Sharing	*	*	*	*	*	*
[1]Vigilance	*	X	X	*	X	X
[1]Visuospatial working memory	*	X		*	*	
[2]Written Comprehension	*	X	*	*	*	*
[2]Written Expression	X	X	*	*	*	*

*Notes. Sources: [1]CASL [2]O*NET. An "X" denotes aptitudes/traits that were documented in the career field background materials through KSAs needed for the career field and/or tasks performed in the career field. An asterisk (*) indicates aptitudes/traits that were not explicitly documented as important for the career field in the archival materials, but were considered important based on the information provided.*

relative importance analysis (see Figure 1) identified these abilities as being relatively more critical than others. If these are also determined to be gaps based on coverage in existing Air Force measures, there is the potential that measuring these abilities in a serious game could provide incremental validity of existing measures.

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19

Of the 19 total traits identified as important for cyber positions, 11 were covered in some capacity in the archival Air Force occupation materials. This left eight traits identified in civilian positions that were not accounted for by the Air Force positions (about 42% of the examined traits). The eight traits not covered in the Air Force archival materials are as follows: Adaptability/Flexibility, Dependability, Independence, Initiative, Persistence, Self-Control, Social Orientation, and Stress Tolerance. Notably, five of the 11 covered traits had little coverage in the occupational materials, namely, Achievement/Effort, Concern for Others, Cooperation, Need for Cognitive Closure, and Tolerance for Risk.

Most of the traits with little to no coverage were identified as highly important through our relative importance computation, especially Adaptability/Flexibility, Dependability, Self-Control, and Stress Tolerance. Moreover, Adaptability/Flexibility, Social Orientation, and Stress Tolerance are particularly suited for measurement through serious games and offer potential incremental validity to current measures if they are confirmed as actual coverage gaps within existing Air Force measures.

Table 6. Summary of Traits Important to Cyber Occupations

Trait	**3D0X2**	**3D0X3**	**1B4X1**	**1N4X1A**	**17DX**	**17SX**
[2]Achievement/Effort			*		*	*
[2]Adaptability/ Flexibility						
[2]Analytical Thinking	*	*	X	*	X	X
[2]Attention to Detail	*	X	X		X	X
[2]Concern for Others					*	*
[2]Cooperation	*					
[2]Dependability						
[2]Independence						
[2]Initiative						
[2]Innovation			X		X	X
[2]Integrity	X	X	*		X	X
[2]Leadership					X	X
[1]Need for Cognition	X	*	X		X	X
[1]Need for Cognitive Closure					*	*
[2]Persistence						
[2]Self Control						
[2]Social Orientation						
[2]Stress Tolerance						
[1]Tolerance for Risk	*	*				

Notes. Sources [1]CASL [2]O*NET. An “X” denotes aptitudes/traits that were documented in the career field background materials through KSAs needed for the career field and/or tasks performed in the career field. An asterisk (*) indicates aptitudes/traits that were not explicitly documented as important for the career field in the archival materials but were considered important based on the information provided.

2.4 Interviews with Subject Matter Experts

Although not explicitly required by the statement of work, we interviewed 11 cyber SMEs to provide further insight into the important aptitudes and traits for the relevant AFSs. We spoke to at least one SME from each AFS, with the exception of the 3D0X2 and 3D0X3 positions. However, an SME who had previously worked in a 3D specialty spoke to our questions based on their experience while in the 3D0X occupations. Additionally, for the 1N4 specialty, we spoke to an SME who had polled other instructors teaching cyber courses for their input as well, thus providing us a wider range of perspectives based upon those additional SMEs. Finally, several SMEs had received training from other agencies (e.g., NSA) and worked in teams comprised of individuals from non-Air Force agencies (those mentioned include: NSA, Navy, and Army). Based on the breadth of experience by the collective group of SMEs that were interviewed, we are confident that coverage of each AFSC is sufficient.

Table 7 below summarizes the aptitudes and traits that SMEs mentioned during the interviews. The source column in the table indicates aptitudes and traits that overlap with the O*NET or CASL detailed above and where the SME generated label means the SME provided a different aptitude or trait that our previous research did not capture. Traits and aptitudes previously discussed that are omitted from Table 7 were not mentioned when SMEs were asked, “What helps to differentiate between the most successful workers versus average ones in terms of abilities, skills, and traits?”

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19

Table 7. SME Generated Aptitudes and Traits

Aptitudes	Source	# of SMEs who endorsed
Achievement/Effort	O*NET	7
Team Player	SME Generated	7
Flexibility/Adaptability	SME Generated	4
Need for Cognition	CASL	4
Self-Discipline	SME Generated	4
Autonomous/Independence	O*NET	3
Creative	SME Generated	3
Oral Expression	O*NET	3
Problem Sensitivity	O*NET	3
Complex Problem Solving	CASL	2
Conscientiousness	O*NET	2
Deductive Reasoning	O*NET	2
Inductive Reasoning	O*NET	2
Instructing/Teaching	SME Generated	2
Openness to experience	SME Generated	2
Oral Comprehension	O*NET	2
Resilience	SME Generated	2
Spatial Reasoning	SME Generated	2
Stress Tolerance	O*NET	2
Tolerance for Ambiguity	SME Generated	2
Written Comprehension	O*NET	2
Written Expression	O*NET	2
Initiative	O*NET	1
Active Learning	SME Generated	1
Category Flexibility	O*NET	1
Cynical view of human nature	SME Generated	1
Decision Making	SME Generated	1
Emotional Intelligence	SME Generated	1
Flexibility of Closure	O*NET	1
Integrity	O*NET	1
Leadership	O*NET	1
Mathematical Reasoning	O*NET	1
Mental Agility	SME Generated	1
Multitasking	SME Generated	1
Number Facility	O*NET	1
Selective Attention	O*NET	1
Situational Awareness	SME Generated	1
Skepticism	SME Generated	1
Speed of Closure	O*NET	1
Systems Thinking	SME Generated	1

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19

Time management	SME Generated	1
Tolerance	SME Generated	1
Visualization	O*NET	1
Confident	SME Generated	1
Memorization	O*NET	1
Mission-oriented	SME Generated	1
Humility	SME Generated	1

Consistent with the O*NET data, multiple SMEs identified Oral Expression and Comprehension, Complex Problem Solving, Inductive and Deductive Reasoning, and Written Expression and Comprehension as important abilities for Air Force cyber positions. SME input demonstrated similar consistency with the O*NET and CASL sourced traits, specifically Achievement/Effort, Need for Cognition, Independence, Problem Sensitivity, Conscientiousness, and Stress Tolerance. SMEs also generated several traits and abilities that were not previously identified. The SME generated traits and aptitudes that were mentioned most frequently were Team Player, Flexibility/Adaptability, Self-Discipline, and Creative. Interestingly, eight out of the top nine constructs of importance identified by SMEs were traits. This underscores the importance of identifying stable characteristics important for success in cyber career fields in addition to other selection information (e.g., certifications, ASVAB scores).
Several of the SME generated traits and abilities overlap with abilities identified as important for cyber security professionals. Specifically, Jose, LaPort, and Trippe (2016) stressed the importance of the abilities related to Working Memory, Cognitive Flexibility, and Systems Thinking. Working memory, or "memorization" as indicated by an SME, is the system that holds multiple pieces of transitory information in the mind, where they can be manipulated (Baddeley & Hitch, 1974)" (Jose et al., 2016, p. 173). Working memory is especially important for cyber-security personnel as they must digest and monitor large quantities of information on a daily basis (Jose et al., 2016). For instance, one of the main responsibilities outlined in the AFECD for 1N4s is to analyze and exploit intelligence information. This involves understanding and acting on large amounts of information. O'Rourke and colleagues (2017) also recognized Working Memory, specifically visuospatial working memory, as an important aptitude for cyber related AFSs.

Cognitive Flexibility, or "Mental Agility" or "Adaptability/Flexibility", refers to "a person's ability to restructure his or her knowledge as an adaptive response to changing situational demands" (Jose et al., 2016, p. 173). Jose and colleagues (2016) highlight the importance of this aptitude for cyber personnel because of the demanding activities the occupations require, such as adapting to new technologies that adversaries may be able to exploit. Cognitive Flexibility relates to two abilities previously discussed, Complex Problem Solving and Category Flexibility.

Systems thinking is another ability considered important for cyber occupations by both SME and literature sources. Systems Thinking is defined as "an approach to problem-solving in which an individual possesses an understanding of how multiple parts of a system interact and influence each other (Aronson, 1996)" (Jose et al., 2016, pp. 173–174). Similar to Complex Problem Solving and Information Ordering, Systems Thinking is important for understanding the interactions between various system components, such as a 1N4s requirement to gain and maintain knowledge of intelligence organizations and systems (AFECD).

As for traits, several SMEs endorsed personal characteristics, such as “active learning”, “thinking outside the box”, and “creative” as important for cyber positions. Jose and colleagues (2016) also identified openness/intellectance as important for cyber professionals. It is imperative that cyber personnel are able to develop solutions to emerging threats and continually learn more about their field. Multiple SMEs emphasized that the people who excel in the cyber career fields are the ones who are passionate about learning more; they “eat, sleep, and breathe this work” and “go home at night and program”.

Thus, through the review of AFSC archival materials and comparison to the integrated abilities and traits sourced from O*NET and CASL, we were able to identify several gaps. These gaps illuminated specific abilities and traits that would be useful to consider for inclusion in serious games used to identify individuals for cyber AFSs. The O*NET data, literature review, and SME input yielded 38 distinct cyber aptitudes and 24 distinct cyber traits. Appendix A provides the complete list of cyber aptitudes and traits identified through the above methods and the corresponding definitions.

2.5 SME Survey

With the identified critical cyber aptitudes and traits, we created a job analysis questionnaire for job incumbents to complete. The questionnaire was composed of three parts; 1) demographic questions, 2) questions about cyber aptitudes, and 3) questions about cyber traits. For Part 2, SMEs were prompted to review a list of 38 cyber aptitudes and their corresponding definitions. After reviewing the list, they were asked to rate each ability on its importance to the performance of their position on a Likert scale (1 = not at all important to 5 = extremely important) and indicate whether each ability was required at entry (yes or no). Then, because prior job analysis research suggests that SMEs tend to inflate importance ratings in job analysis questionnaires (Gael, 1988; Morgeson & Campion, 1997) they were also prompted to select the top five (most important) abilities and the bottom five (least important) abilities of the 38 to further differentiate the most critical abilities. Participants repeated this process for the 24 cyber traits in Part 3. Participants who were identified as SMEs in cyber positions were forwarded the survey and given the opportunity to participate in the research. Participants could complete the survey through Qualtrics online or through a PDF paper version. Prospective participants were informed that the goal of the data collection effort was to identify critical cyber abilities and traits for cyber positions. The recruitment email sent to SMEs is reproduced in Appendix B. The three survey parts can be found in Appendix C.

Sixty-one SMEs completed the online questionnaire through Qualtrics. All participants were employed by the Air Force at the time of the survey. Sixteen of the participants only completed the demographics (Part 1) and were not included in the final sample. The final sample ($N = 45$) was 91.3% male, with a mean age of 30.57 years ($SD = 5.89$). Participants had worked in their current position between 1 month and 19 years ($M = 2.49$, $SD = .12$). Most participants had at least one cyber-related certification (84.8%), and a large percentage of participants categorized their educational level as associate’s degree (34.8%) or bachelor’s degree (30.4%), whereas

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19

smaller numbers reported the following categories: some college courses (15.2%), high school diploma/GED (10.9%), or post-graduate degree (8.7%). In order to evaluate the survey results by AFSC, participants were grouped by a general AFSC of 3Ds, 1Bs, 1Ns, or 17s (see Table 8).

Table 8. SME Survey Participant AFSCs

AFSC – General	AFSC – Specific	Frequency	%
3D	3D032	1	6%
	3D052	3	18%
	3D072	3	18%
	3D074	1	6%
	3D0X2	3	18%
	3D151	1	6%
	3D152	1	6%
	3D171	1	6%
	3D172	1	6%
	3D173	2	12%
	Total	**17**	**100%**
1B	1B431	1	5%
	1B451	2	10%
	1B471	12	57%
	1B491	2	10%
	1B4X1	4	19%
	Total	**21**	**100%**
1N	1N471A	2	40%
	1N490	1	20%
	1N4A	2	40%
	Total	**5**	**100%**
17	17D	1	50%
	17S3D	1	50%
	Total	**2**	**100%**

Attention to detail, analytical thinking, active learning, stress tolerance, and adaptability were rated highly across AFSCs with high agreement among SME raters. For the AFSCs with acceptable sample sizes, there were few average importance ratings below 3 (Moderately Important). This is consistent with previous job analytic findings where importance ratings are inflated (Gael, 1988; Morgeson & Campion, 1997). There were no aptitudes or traits rated as less than 3 for the 3Ds. For the 1Bs, emotional intelligence, fluency of ideas, instructing, mathematical reasoning, and social orientation were all rated below 3, suggested these aptitudes and trait are relatively less important. All average importance ratings by AFSC are reported in Table 9 though the 1N and 17 results should be interpreted with caution as there were fewer than 10 SMEs in each of those AFSCs.

We used r_{wg} to assess the level of interrater agreement among SME raters for each aptitude/trait within each AFSC. This index reflects the extent to which individuals agree on the rating of a stimulus compared to the level of agreement that would be expected by chance (James, Demaree, & Wolf, 1984). R_{wgs} were not calculated for 1Ns and 17s because there were too few raters. Fewer than 10 raters can result in the attenuation of r_{wgs} (LeBreton & Senter, 2008). Still, the

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19

1Ns and 17s were included in the calculation of the overall r_{wgs}. According to LeBreton and Senter (2008), r_{wgs} should be at least .50 to suggest moderate agreement, while r_{wgs} greater than .70 suggest strong agreement. There was weak agreement on the importance of fourteen of the 62 aptitudes/traits for the 3Ds and for 24 of the 62 aptitudes/traits for the 1Bs. (see Table 9). The remaining aptitudes/traits had at least moderate agreement for the importance ratings (greater than .50).

Table 9. SME Average Importance Ratings and r_{wg} by AFSC

	3Ds N=16-17		**1Bs** N=15-20		***1Ns*** *N=4-5*		***17s*** *N=1-2*		**Overall** N=37-45	
	Mean	r_{wg}	Mean	r_{wg}	*Mean*	r_{wg}	*Mean*	r_{wg}	Mean	r_{wg}
Attention to detail	4.88	0.94	4.80	0.91	5.00	N/A	4.00	N/A	4.84	0.93
Analytical thinking	4.75	0.90	4.80	0.91	5.00	N/A	5.00	N/A	4.81	0.92
Information and technology aptitude	4.82	0.86	4.45	0.55	4.80	N/A	4.50	N/A	4.64	0.72
Active learning	4.71	0.89	4.60	0.82	5.00	N/A	4.00	N/A	4.62	0.83
Stress tolerance	4.69	0.82	4.53	0.87	4.25	N/A	5.00	N/A	4.59	0.82
Adaptability	4.75	0.90	4.33	0.81	4.25	N/A	4.00	N/A	4.51	0.82
Initiative	4.56	0.74	4.60	0.80	4.00	N/A	5.00	N/A	4.49	0.73
Persistence	4.50	0.80	4.67	0.88	4.00	N/A	4.00	N/A	4.46	0.79
Resilience	4.63	0.81	4.53	0.72	3.75	N/A	5.00	N/A	4.46	0.71
Deductive reasoning	4.59	0.87	4.55	0.82	3.60	N/A	5.00	N/A	4.44	0.78
Integrity	4.44	0.27	4.40	0.51	4.25	N/A	5.00	N/A	4.43	0.48
Written comprehension	4.47	0.81	4.35	0.51	4.80	N/A	5.00	N/A	4.43	0.67
Complex problem-solving	4.71	0.89	4.45	0.50	3.80	N/A	4.00	N/A	4.42	0.60
Dependability	4.75	0.90	4.07	0.47	4.25	N/A	5.00	N/A	4.41	0.68
Situational awareness	4.63	0.74	4.47	0.72	4.50	N/A	2.00	N/A	4.41	0.60
Time management	4.53	0.68	4.30	0.73	4.20	N/A	4.00	N/A	4.39	0.67
Decision making	4.35	0.38	4.35	0.67	4.00	N/A	5.00	N/A	4.36	0.56
Vigilance	4.12	0.32	4.45	0.82	4.40	N/A	4.00	N/A	4.32	0.61
Systems thinking	4.53	0.74	4.25	0.69	3.60	N/A	5.00	N/A	4.27	0.67
Self-discipline	4.38	0.68	4.20	0.63	4.50	N/A	3.00	N/A	4.27	0.68
Oral comprehension	4.41	0.75	4.00	0.53	4.20	N/A	5.00	N/A	4.23	0.63
Self control	4.31	0.62	4.20	0.63	4.00	N/A	4.00	N/A	4.22	0.66
Team player	4.25	0.63	3.87	0.65	5.00	N/A	4.00	N/A	4.19	0.64
Mental agility	4.29	0.76	4.25	0.69	4.00	N/A	3.00	N/A	4.16	0.61
Need for cognition	4.19	0.65	4.20	0.63	3.75	N/A	5.00	N/A	4.14	0.63
Inductive reasoning	4.24	0.72	4.20	0.71	3.80	N/A	3.00	N/A	4.09	0.71
Cooperation	4.31	0.62	3.93	0.90	3.50	N/A	4.00	N/A	4.05	0.75
Pattern recognition and scanning	4.35	0.63	4.00	0.42	3.60	N/A	3.00	N/A	4.05	0.51
Convergent creative thinking	4.06	0.72	4.00	0.47	4.00	N/A	4.00	N/A	4.04	0.61
Independence	4.38	0.61	3.80	0.63	4.25	N/A	4.00	N/A	4.03	0.51
Problem sensitivity	4.00	0.44	4.15	0.62	3.80	N/A	4.00	N/A	4.02	0.59
Anomaly detection	4.24	0.53	4.21	0.47	3.40	N/A	2.00	N/A	4.00	0.42
Innovation	4.06	0.57	3.80	0.41	4.50	N/A	5.00	N/A	4.00	0.50
Written expression	4.12	0.51	3.60	0.03	4.80	N/A	5.00	N/A	3.95	0.28
Oral expression	4.12	0.63	3.45	0.19	4.60	N/A	5.00	N/A	3.91	0.35
Selective attention	3.65	0.57	4.10	0.64	4.00	N/A	3.00	N/A	3.91	0.56
Tolerance for risk	4.13	0.68	3.80	0.34	3.00	N/A	2.00	N/A	3.84	0.46
Time sharing	3.65	0.57	3.90	0.32	4.20	N/A	4.00	N/A	3.82	0.48

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19

Need for cognitive closure	4.06	0.57	3.80	0.56	2.75	N/A	5.00	N/A	3.81	0.45
Achievement	3.75	0.37	3.80	0.56	4.00	N/A	4.00	N/A	3.78	0.52
Category flexibility	4.06	0.72	3.60	0.61	3.80	N/A	2.50	N/A	3.73	0.65
Memorization	3.59	0.56	3.80	0.55	4.20	N/A	2.00	N/A	3.71	0.51
Information ordering	4.18	0.55	3.60	0.61	3.20	N/A	2.00	N/A	3.69	0.50
Modeling program execution	3.76	0.65	3.90	0.11	3.20	N/A	2.50	N/A	3.69	0.39
Speed of closure	3.71	0.64	3.75	0.32	3.40	N/A	4.00	N/A	3.68	0.47
Perceptual speed	3.94	0.53	3.75	0.64	3.00	N/A	1.00	N/A	3.66	0.51
Mental model ability	3.71	0.70	3.65	0.35	3.60	N/A	3.00	N/A	3.62	0.54
Resistance to interference	3.71	0.58	3.70	0.68	3.60	N/A	1.00	N/A	3.61	0.58
Psychomotor speed	3.65	0.82	3.70	0.68	3.00	N/A	1.00	N/A	3.52	0.62
Leadership	4.00	0.47	3.00	0.07	3.50	N/A	4.00	N/A	3.51	0.23
Flexibility of closure	3.71	0.70	3.65	0.25	2.80	N/A	2.00	N/A	3.49	0.42
Visuospatial working memory	3.65	0.69	3.35	0.62	3.40	N/A	1.00	N/A	3.41	0.60
Fluency of ideas	4.00	0.44	2.85	0.41	3.20	N/A	3.00	N/A	3.33	0.27
Spatial visualization	3.47	0.37	3.15	0.46	3.40	N/A	1.00	N/A	3.25	0.35
Skepticism	3.81	0.39	3.00	0.36	2.50	N/A	2.00	N/A	3.24	0.29
Number facility	3.41	0.56	3.10	0.22	3.00	N/A	2.00	N/A	3.18	0.39
Instructing	3.65	-0.06	2.65	0.35	3.60	N/A	4.50	N/A	3.18	0.04
Social orientation	3.44	0.34	2.87	0.22	3.25	N/A	3.00	N/A	3.14	0.30
Concern for others	3.06	0.04	3.20	0.49	2.75	N/A	2.00	N/A	3.08	0.24
Originality	3.18	0.36	3.00	0.16	3.00	N/A	4.00	N/A	3.05	0.28
Emotional intelligence	3.29	0.39	2.75	0.06	2.80	N/A	3.00	N/A	2.98	0.22
Mathematical reasoning	3.59	0.56	2.55	0.29	2.80	N/A	2.00	N/A	2.98	0.28

The top five ratings differentiated between the important aptitudes and traits more effectively than the Likert scale ratings. The results of the top five ratings indicated by the percentage in the top five for each AFSC are in Table 10 (for the aptitudes) and 11 (for the traits).

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19

Table 10. SME Top Five Percentages by AFSC for Aptitudes

Tier	Aptitude	3Ds *N=17*	1Bs *N=20*	1Ns *N=5*	17s *N=2*
Tier I	Decision making	41.2%	60.0%	20.0%	100.0%
	Complex problem-solving	52.9%	55.0%	20.0%	50.0%
	Active learning	88.2%	50.0%	60.0%	0.0%
	Deductive reasoning	41.2%	40.0%	20.0%	50.0%
	Information and technology aptitude	47.1%	35.0%	40.0%	0.0%
	Anomaly detection	41.2%	35.0%	0.0%	0.0%
Tier II	Written comprehension	11.8%	30.0%	40.0%	0.0%
	Time management	23.5%	20.0%	20.0%	0.0%
	Inductive reasoning	5.9%	20.0%	20.0%	0.0%
	Mental agility	5.9%	20.0%	40.0%	0.0%
	Systems thinking	23.5%	15.0%	0.0%	50.0%
	Oral expression	23.5%	0.0%	40.0%	50.0%
Tier III	Convergent creative thinking	11.8%	15.0%	20.0%	0.0%
	Pattern recognition and scanning	5.9%	15.0%	20.0%	0.0%
	Problem sensitivity	0.0%	15.0%	0.0%	0.0%
	Written expression	11.8%	10.0%	0.0%	0.0%
	Emotional intelligence	0.0%	10.0%	0.0%	50.0%
	Number facility	0.0%	10.0%	0.0%	0.0%
	Instructing	17.6%	5.0%	0.0%	50.0%
	Information ordering	11.8%	5.0%	0.0%	0.0%
	Vigilance	11.8%	5.0%	0.0%	0.0%
	Memorization	5.9%	5.0%	0.0%	0.0%
	Oral comprehension	5.9%	5.0%	20.0%	0.0%
	Psychomotor speed	0.0%	5.0%	0.0%	0.0%
	Selective attention	0.0%	5.0%	0.0%	0.0%
	Speed of closure	0.0%	5.0%	0.0%	0.0%
	Time sharing	0.0%	5.0%	20.0%	0.0%
	Visuospatial working memory	0.0%	5.0%	0.0%	0.0%
	Mathematical reasoning	5.9%	0.0%	0.0%	0.0%
	Modeling program execution	5.9%	0.0%	0.0%	0.0%
	Originality	5.9%	0.0%	0.0%	0.0%
	Category flexibility	0.0%	0.0%	0.0%	0.0%
	Flexibility of closure	0.0%	0.0%	0.0%	0.0%
	Fluency of ideas	0.0%	0.0%	0.0%	0.0%
	Mental model ability	0.0%	0.0%	0.0%	0.0%
	Perceptual speed	0.0%	0.0%	0.0%	0.0%
	Resistance to interference	0.0%	0.0%	0.0%	0.0%
	Spatial visualization	0.0%	0.0%	0.0%	0.0%

Note: Tier I = 3D or 1B4X1 is above 30%; Tier II = 3D and/or 1B4X1 are equal to or between 20% and 30%; Tier III = 3D and 1B4X1 are below 20%.

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19

Table 11. SME Top Five Percentages by AFSC for Traits

Tier	Trait	3Ds *N=16*	1B4X1 *N=15*	1Ns *N=4*	17s *N=1*
Tier I	Analytical thinking	68.8%	66.7%	25.0%	100.0%
	Attention to detail	68.8%	60.0%	75.0%	0.0%
	Adaptability	50.0%	33.3%	50.0%	100.0%
	Initiative	62.5%	26.7%	75.0%	0.0%
	Dependability	56.3%	20.0%	25.0%	0.0%
	Persistence	25.0%	33.3%	25.0%	0.0%
	Integrity	18.8%	33.3%	0.0%	0.0%
	Stress tolerance	12.5%	46.7%	25.0%	0.0%
Tier II	Team player	25.0%	13.3%	25.0%	0.0%
	Self-discipline	12.5%	26.7%	0.0%	0.0%
	Situational awareness	12.5%	26.7%	25.0%	0.0%
	Resilience	0.0%	20.0%	25.0%	0.0%
Tier III	Innovation	12.5%	13.3%	0.0%	100.0%
	Leadership	12.5%	13.3%	25.0%	0.0%
	Achievement	12.5%	6.7%	0.0%	0.0%
	Tolerance for risk	12.5%	6.7%	0.0%	0.0%
	Cooperation	12.5%	0.0%	0.0%	0.0%
	Need for cognition	6.3%	13.3%	0.0%	100.0%
	Self control	6.3%	13.3%	25.0%	0.0%
	Concern for others	6.3%	6.7%	25.0%	0.0%
	Independence	6.3%	6.7%	25.0%	0.0%
	Skepticism	6.3%	0.0%	0.0%	0.0%
	Need for cognitive closure	0.0%	13.3%	0.0%	100.0%
	Social orientation	0.0%	6.7%	25.0%	0.0%

Note: Tier I = 3D or 1B4X1 is above 30%; Tier II = 3D and/or 1B4X1 are equal to or between 20% and 30%; Tier III = 3D and 1B4X1 are below 20%.

Ultimately, we computed a weighted average based on the sample size for each AFSC and percentage who indicated the aptitude/trait should be in the top five to identify the top 10 most critical traits (see Tables 12 and 13).

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19

Table 12. Ten Most Critical Aptitudes

Aptitude	3Ds %Top 5	3Ds N	1B4X1 %Top 5	1B4X1 N	1Ns %Top 5	1Ns N	17s %Top 5	17s N	Weighted average
Active learning	88%	17	50%	20	60%	5	0%	2	**63.6%**
Decision making	41%	17	60%	20	20%	5	100%	2	**50.0%**
Complex problem-solving	53%	17	55%	20	20%	5	50%	2	**50.0%**
Information and technology aptitude	47%	17	35%	20	40%	5	0%	2	**38.7%**
Deductive reasoning	41%	17	40%	20	20%	5	50%	2	**38.6%**
Anomaly detection	41%	17	35%	20	0%	5	0%	2	**31.8%**
Written comprehension	12%	17	30%	20	40%	5	0%	2	**22.7%**
Time management	24%	17	20%	20	20%	5	0%	2	**20.4%**
Systems thinking	24%	17	15%	20	0%	5	50%	2	**18.2%**
Mental agility	6%	17	20%	20	40%	5	0%	2	**15.9%**

Table 13. Ten Most Critical Traits

Trait	3Ds %Top 5	3Ds N	1B4X1 %Top 5	1B4X1 N	1Ns %Top 5	1Ns N	17s %Top 5	17s N	Weighted average
Analytical thinking	69%	16	67%	15	25%	4	100%	1	**63.9%**
Attention to detail	69%	16	60%	15	75%	4	0%	1	**63.9%**
Initiative	63%	16	27%	15	75%	4	0%	1	**47.2%**
Adaptability	50%	16	33%	15	50%	4	100%	1	**44.4%**
Dependability	56%	16	20%	15	25%	4	0%	1	**36.1%**
Stress tolerance	13%	16	47%	15	25%	4	0%	1	**27.8%**
Persistence	25%	16	33%	15	25%	4	0%	1	**27.8%**
Integrity	19%	16	33%	15	0%	4	0%	1	**22.2%**
Team player	25%	16	13%	15	25%	4	0%	1	**19.4%**
Situational awareness	13%	16	27%	15	25%	4	0%	1	**19.4%**

Still, the importance ratings suggest that SMEs viewed all the aptitudes/traits as at least somewhat important. Thus, to the extent possible, the following sections consider all the aptitudes/traits identified as important through the literature review and SME interviews.

2.6 Differences between Officer and Enlisted Cyber Career Fields

The issue of if there are different requirements for officers and enlisted working in the cyber career fields is an important one. We addressed this question through our SMEs by asking them if there were significant differences between enlisted and officer positions. By and large, the consensus was that officers needed to know the general problem area so they could interact meaningfully with their teams but did not need to know specific technical information (e.g., how

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19

to program a router). Officers could best support their teams by watching out for and developing their personnel, behaviors that one SME termed “traditional officer behaviors”. Thus, effective leadership is an important tool cyber officers need to have; and the consensus from SMEs is that the Air Force already does a good job training the leadership skills needed by officers.

One area that may require special attention for officers is the aptitudes and traits relevant to the communication and influence domains. Officers are often required to clearly describe current capabilities, requirements, and shortfalls to ensure organizations properly plan and direct cyber operations. Given the technical information and terminology that accompany cyber warfare and network operations, officers that are able to succinctly interpret and summarize relevant information into meaningful updates and actionable intelligence may prove uniquely valuable to commanders and staffs. Further exploration into this idea is needed to empirically confirm these qualitative assumptions regarding the unique importance of superior communication skills for leaders within the cyber domain.

2.7 Summary

Our first objective was to review the military and civilian literatures to identify those aptitudes and traits important to cyber career fields in the Air Force. These literatures proved extensive and informative. We also asked cyber SMEs to generate aptitudes and traits necessary for competency in cyber AFSCs. Given these three sources of input, the SMEs provided a rank ordering based on importance. The top ten aptitudes and traits are provided above in Tables 12 and 13, respectively. With the aptitudes and traits necessary for competency in cyber occupations identified, we move to objective two which focused on assessing the utility of existing measures of those aptitudes and traits.

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19

3.0 OBJECTIVE 2 – EXISTING MEASURES OF APTITUDES & TRAITS

The military and civilian worlds each provide assessments of various traits and aptitudes of interest to us. Our focus in this section is to consider how well the criterion space of the constructs of interest is covered by existing measures for each of the aptitudes and traits considered in objective 1. We conclude the section with a gap analysis pointing out where assessments important traits and aptitudes for cyber performance are lacking or need further development.

3.1 Best bets for measuring cyber aptitudes and traits using existing DoD tests

We reviewed several existing DoD tests, which are identified below. The reviewed tests include those used for both enlisted and officer positions. Our review included tests administered to enlisted airmen, namely the Armed Services Vocational Aptitude Battery (ASVAB) and Tailored Adaptive Personality Assessment System (TAPAS); as well as tests administered to officers, specifically the Air Force Officer Qualifying Test (AFOQT) and the Self-Description Inventory (SDI). We also covered several domain-specific tests, including the: Air Force Cyber Test, Electronic Data Processing Test (EDPT), Cyber Aptitude and Talent Assessment (CATA) battery, and the Air Force Multitasking Test. We next provide information on these assessments and discuss their utility for measuring cyber aptitudes and traits.

3.1.1. Armed Services Vocational Aptitude Battery (ASVAB)

The Air Force began using the ASVAB in 1973 for both selection and classification into enlisted positions (Thompson, 2007). The ASVAB consists of nine subtests: Arithmetic Reasoning (AR), Assembling Objects (AO), Automotive and Shop (AS), Electronics Information (EI), General Science (GS), Mathematics Knowledge (MK), Mechanical Comprehension (MC), Paragraph Comprehension (PC) and Word Knowledge (WK) (see Table 14). Different combinations of the nine subsets are used to compute the Air Force MAGE composites (Mechanical, Administrative, General, and Electronics)[1].

Table 14. ASVAB Subtests and Definitions

ASVAB Subtest	Definition
Arithmetic Reasoning (AR)	Ability to solve arithmetic word problems
Assembling Objects (AO)	Ability to determine how an object will look when its parts are put together
Automotive & Shop (AS)	Knowledge of automobile technology and tools and shop terminology and practices
Electronics Information (EI)	Knowledge of electricity and electronics
General Science (GS)	Knowledge of physical and biological sciences
Mathematics Knowledge (MK)	Knowledge of high school mathematics principles
Mechanical Comprehension (MC)	Knowledge of mechanical and physical principles
Paragraph Comprehension (PC)	Ability to obtain information from written passages
Word Knowledge (WK)	Ability to select the correct meaning of a word presented in context and to identify best synonym for a given word

[1] Additional Information on the ASVAB is available at http://official-asvab.com/understand_coun.htm

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19

While there is a long history of the relatively large predictive validity of cognitive ability measures for job performance (r = .53; Hunter & Hunter, 1984), there is evidence of group differences that favor majority groups in cognitive ability testing, which can lead to disproportionately hiring majority group members when cognitive ability measures are used for selection decisions (Roth, Bevier, Bobko, Switzer, & Tyler, 2001). The ASVAB showed these same patterns with high predictive validity for job performance (ranging from .29 to .87 with a median of .68; Thompson, 2007) (see Fairbanks, Kucinkas, Nakasone, Trent, & Welsh, 1989 for a review of ASVAB validity studies) and substantial group differences (Haywood, Palmer, & Ward, 1990; Looper, 1997; Ree, Valentine, & Wilbourn, 1984; Welsh, 1997).

ASVAB scores are used for qualification into several of the enlisted cyber positions covered in the present research. The minimum ASVAB score required for entry into 3D0X2 (Cyber Systems Operations), 3D0X3 (Cyber Surety), and 1B4X1 (Cyber Warfare Operations) specialties is a General (G) composite score (Arithmetic Reasoning [AR] + Verbal Expression [VE]) of 64 (or a lower ASVAB G composite score combined with a score of 60+ on the Cyber Test) (Air Force Enlisted Classification Directory (AFECD), 2017). The 1N4X1A (Digital Network Analyst) specialty requires a slightly lower minimum ASVAB G composite score of 62 (or a lower ASVAB G score combined with a score of 60+ on the Cyber Test) (Air Force Enlisted Classification Directory (AFECD), 2017).

The Air Force Personnel Center (AFPC) evaluated the ASVAB entry standards with a sample of 1N4X1A airmen and provided the results in a briefing (Evaluation of 1N4X1A/B Entry Standards, 2017). They examined the validity of ASVAB composites and subtests for predicting completion of the Joint Cyber Analysis Course (JCAC). The JCAC training requirement was identified as a significant choke point leading to an attrition rate over 36% in between 2014 to 2016 for Digital Network Analysts (1N4X1A) trainees. The Apprentice Course (AC) also had notable attrition rates from 2014 to 2016 (8.82% for 1N4X1A).

According the AFPC's evaluation of 1N4X1A entry standards, the General MAGE composite is useful for predicting JCAC success (adjusted R^2 = .14, r = .37) in a sample of 253 1N4X1A airmen. Including scores from the GS and MK subtests as predictors adds incremental predictive validity (adjusted R^2 = .17; ΔR^2 = .03). The AR, EI and PC subtests were also listed in this study as being significantly positively correlated with JCAC of JCAC course completion (r's = .19 to .23) but did not retain significance in follow up regression analyses when MAGE-G, GS, MK, and Tolerance were taken into account. The remaining subtests (AO, AS, MC, and WK) were not listed as being significantly correlated with JCAC success.

3.1.2. Tailored Adaptive Personality Assessment System (TAPAS)

TAPAS is used to augment the predictive power of the ASVAB for personnel selection and classification decisions for enlisted airmen. TAPAS was developed to assess several facets of the Big Five personality factors as well as additional personality characteristics relevant to military settings (see Tables 15 and 16). Of particular importance is that TAPAS is designed to be resistant to faking, so it can be used for high stakes assessment as found in enlistment testing. Each TAPAS item consists of two statements, balanced in social desirability, and a respondent picks the statement that is "more like me."

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19

Unlike cognitive ability tests, personality tests typically have little, if any, adverse impact against minority groups (Foldes, Duehr, & Ones, 2008). Early results from an initial operational test and assessment indicated little adverse impact for females and ethnic minority groups (Drasgow et al., 2012) as well as good resistance to faking (Stark, Chernyshenko, Nye, Drasgow, & White, 2017). Though much of the validation work for TAPAS is ongoing (Stark, Chernyshenko, & Drasgow, 2012), meta-analyses demonstrate consistent incremental validity of personality traits over cognitive ability (Barrick & Mount, 1991; Schmidt & Hunter, 1998).

Traits currently measured by the US Air Force using TAPAS were identified from the document "AF Tailored Adaptive Personality Assessment System - CY 18.docx" provided by the Air Force. This document lists the fifteen subscales currently being used or that will be implemented in the near future: Achievement, Adjustment, Attention-Seeking, Cooperation, Dominance, Even-Tempered, Non-Delinquency/Traditionalism, Optimism, Physical Condition, Responsibility, Self-Control, Selflessness, Situational Awareness, Sociability, and Tolerance. It also includes a description of behaviors representative of high-scoring individuals for each scale.

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19

Table 15. TAPAS Traits and Definitions

Trait	Description *High-scoring individuals…*
Achievement*	are seen as hard working, ambitious, confident, and resourceful.
Adjustment*	are worry free, and handle stress well.
Attention Seeking*	tend to engage in behaviors that attract social attention; they are loud, talkative, entertaining, and even boastful.
Adventure Seeking	enjoy participating in extreme sports and outdoor activities.
Aesthetics	appreciate various forms of art and music and participate in art-related activities more often than others.
Attention-Seeking	tend to engage in behaviors that attract social attention; they are loud, entertaining, and even boastful.
Consideration	are affectionate, compassionate, sensitive, and caring.
Cooperation*	are trusting, cordial, non-critical, and easy to get along with.
Courage	stand up to challenges and are not afraid to face dangerous situations
Curiosity	are inquisitive and perceptive, they are interested in learning new information and attend courses and workshops whenever they can.
Depth	exhibit behaviors targeted toward understanding the meaning of one's life and/or facilitating self-improvement, reflection, and self-actualization.
Dominance*	are domineering, take charge, and are often referred to by their peers as "natural leaders."
Even-Tempered*	tend to be calm and stable. They don't often exhibit anger, hostility, or aggression.
Ingenuity	are inventive and can think "outside of the box."
Intellectual Efficiency	are able to process information quickly and would be described by others as knowledgeable, astute, and intellectual.
Non-Delinquency*	tend to comply with rules, customs, norms, and expectations, and they tend not to challenge authority.
Optimism*	have a positive outlook on life and tend to experience joy and a sense of well-being.
Order	tend to organize tasks and activities and desire to maintain neat and clean surroundings.
Physical Conditioning*	tend to engage in activities to maintain their physical fitness and are more likely to participate in vigorous sports or exercise.
Responsibility*	are dependable, reliable, and make every effort to keep their promises.
Self-Control*	tend to be cautious, levelheaded, able to delay gratification, and patient.
Selflessness/Generosity*	are generous with their time and resources.
Situational Awareness*	pay attention to their surroundings and rarely get lost or surprised.
Sociability*	tend to seek out and initiate social interactions.
Team Orientation	prefer working in teams and help people work together better.
Tolerance*	scoring are interested in other cultures and opinions that may differ from their own. They are willing to adapt to novel environments and situations.
Virtue	strive to adhere to standards of honesty, morality, and "good Samaritan" behavior.

Note. Traits indicated with an asterisk (*) are currently measured by the AF TAPAS or scheduled to appear on the next version to be implemented. Descriptions are from Stark et al. (2014) and Nye et al. (2014).

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19

Table 16. TAPAS Mapped onto Big Five Personality Factors

TAPAS Dimension	Big Five Factor O	C	ES	A	E	TAPAS Specific
Aesthetics	X					
Curiosity	X					
Depth	X					
Ingenuity	X					
Intellectual Efficiency	X					
Tolerance	X					
Achievement		X				
Non-Delinquency		X				
Order		X				
Responsibility		X				
Self-Control		X				
Virtue		X				
Adjustment			X			
Attention-Seeking			X			
Even-Tempered			X			
Optimism			X			
Consideration				X		
Cooperation				X		
Selflessness/Generosity				X		
Dominance					X	
Sociability					X	
Adventure Seeking						X
Courage						X
Physical Conditioning*						
Situational Awareness						X
Team Orientation						X

Notes. O = Openness to Experience, C = Conscientiousness, ES = Emotional Stability, A = Agreeableness, E = Extraversion. Based on information from Stark et al. (2014) and Nye et al. (2014).
*There is a lack of consensus on the mapping of the TAPAS Physical Conditioning dimension.

The AFPC has evaluated the predictive validity of TAPAS in combination with other tests as a potential predictor of technical training and JCAC attrition (Evaluation of 1N4X1A/B Entry Standards, 2016). Results from a sample of 99 Digital Network Analysts (1N4X1A) students indicated that combining scores from the GS and MK subtests, without the General MAGE composite as a predictor, with scores on Tolerance disposition (individuals scoring high on this facet are interested in other cultures and opinions that may differ from their own) increased the predictive validity by (ΔR^2 = .07), accounting for 20% of the total variance in course success rates. No other TAPAS dimensions significantly correlated with course success.

3.1.3. Air Force Officer Qualifying Test (AFOQT)

The AFOQT is one of two assessments exclusively administered to candidates for training as officers. The AFOQT is the cognitive ability assessment (Air Force Officer Qualifying Test (AFOQT) Information Pamphlet, 2015) and the Self-Description Inventory (SDI) is the personality assessment. The measured aptitudes are used for several purposes, including: to select candidates into officer commissioning and specific officer training programs, identify

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19

recipients for scholarship awards to the United States Air Force (USAF) Reserve Officer Training Corps, and select applicants for officer commissioning through the ROTC and Officer Training School programs. In conjunction with other occupation-specific requirements, AFOQT scores are used for qualification into aircrew training specialties such as pilots, combat systems operator, air battle manager, and remotely-piloted aircraft pilots.

There have been several revisions to the AFOQT in order to ensure the test content is relevant to officer positions, the assessment is reliable and valid, and the test meets industry standards for adverse impact (Agee, Shore, Alley, Barto, & Halper, 2009). Eighteen versions of the AFOQT were published by the Air Force Human Resources Laboratory between 1951 and 1999 (Alley et al., 2007). The current version of the AFOQT (Form T) contains 10 cognitive subtests that are combined into 6 composites that are used for officer commissioning and aircrew training qualification. The 10 cognitive subtests are Verbal Analogies (VA), Word Knowledge (WK), Reading Comprehension (RC), Arithmetic Reasoning (AR), Math Knowledge (MK), Block Counting (BC), Instrument Comprehension (IC), Aviation Information (AI), Physical Science (PS), and Table Reading (TR). Carretta, Rose, and Trent (2016) provided descriptions of each of the cognitive subtests, which are reproduced in Table 17. These subtests are combined into six composites: Academic, Verbal, Quantitative, Pilot, Air Battle Manager (ABM), and Combat Systems Officer (CSO). The subtests contributions to the composites is shown in Table 18.

Table 17. AFOQT Subtests and Definitions (from Carretta et al., 2016)

Subtest	**Description**
Verbal Analogies	assesses the ability to reason and determine the relations between words.
Arithmetic Reasoning	uses word problems to assess the ability to understand arithmetic relations.
Reading Comprehension	assesses the ability to read and understand written material.
Word Knowledge	measures verbal comprehension of written language involving the use of synonyms.
Math Knowledge	assesses the ability to use mathematical formulas, relations, and terms.
Instrument Comprehension	measures the ability to determine the altitude of an aircraft from illustrations of flight instruments.
Block Counting	provides a measure of spatial ability through the analysis of three-dimensional representation of a set of blocks.
Table Reading	measures the ability to quickly and accurately extract information from tables.
Aviation Information	assesses knowledge of general aviation concepts, principles, and terms.
Physical Science	provides a measure of knowledge and understanding of scientific, terms, concepts, instruments, and principles.

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19

Table 18. Construction of AFOQT Form T Composites

		Composite					
Subtest	N Items	Pilot	CSO	ABM	Academic Aptitude	Verbal	Quant.
Verbal Analogies	25			X	X	X	
Arithmetic Reasoning	25				X		X
Word Knowledge	25		X		X	X	
Math Knowledge	25	X	X	X	X		X
Reading Comprehension	25				X	X	
Physical Science	20						
Table Reading	40	X	X	X			
Instrument Comprehension	25	X		X			
Block Counting	30		X	X			
Aviation Information	20	X		X			

Notes. Physical Science (PS) does not contribute to any of the AFOQT Form T composites. Adapted from (Carretta et al., 2016).

In addition to the cognitive subtests listed above, AFOQT Form T includes a Situational Judgement subtest, which "measures judgment and decision-making in responding to the types of interpersonal situations often encountered by junior USAF officers" (Officer Qualifying Test (AFOQT) Information Pamphlet, 2015). The 50 Situational Judgment questions are based on scenarios inspired by real situations encountered by junior officers that required the use of the core competencies of Integrity and Professionalism, Leadership, Resource Management, Communication, Innovation, and Mentoring ("Officer Qualifying Test (AFOQT) Information Pamphlet," 2015). The situational judgment test (SJT) questions are scored relative to the consensus judgment provided by high-potential USAF officers and test questions were selected for inclusion in the AFOQT based on statistical relationships of scores to cadet outcomes in Basic Officer Training and Field Training (Officer Qualifying Test (AFOQT) Information Pamphlet, 2015). This subtest may map onto several of the aptitudes and traits identified as important for cyber positions. Specifically, the subtest may map onto aptitudes and traits such as Integrity, Leadership, Innovation, and Instructing identified in Objective 1.

Item, test, factor, and composition level analyses for two form versions of the AFOQT [Form T1 (N = 5,681) or Form T2 (N = 5,199)] indicated that the assessment is psychometrically sound (Carretta et al., 2016). Confirmatory factor analyses employing a model with a hierarchical general factor and five first-order factors (verbal, math, spatial, perceptual speed, and aviation knowledge) provided the best fit. Internal consistency was similar for both forms. Cronbach's alpha ranged from .730 (RC) to .913 (IC) for Form T1 and from .741 (AI) to .904 (IC) for Form T2 with mean reliabilities of .816 and .815, respectively. Internal consistency reliability was .80 or higher for six subtests (AR, WK, MK, TR, IC, and BC) on both forms.

There have been several studies to investigate the validity of the AFOQT for predicting officer performance criteria, such as training outcomes, classroom grades, and so forth. Collectively, these studies demonstrate the validity of the AFOQT for officer selection (Alley et al., 2007; Skinner, 2006). Similar to other cognitive ability assessments, studies have found subgroup differences in mean test scores for minority and majority groups (Skinner, 2006).

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19

In the context of cyber training specialties, Manley (2016) examined the predictive value of the AFOQT (Form S) for training success. Training success was operationalized in this study as the final school grade in the 17D Undergraduate Cyber Training Phase II course in the Air Force Officer Cyberspace Operations / Cyberspace Control School (Manley, 2016). In a sample of 17D officers, the researchers found that all AFOQT composites and subtests were significantly related to training success. The uncorrected correlations for each AFOQT composite and subtest are reproduced in Table 19.

Table 19. Prediction of 17D Training Success with AFOQT

Composite	*r*	Subtest	*r*
Pilot	.36***	Verbal Analogies	.33***
Nav/Tech	.35***	Arithmetic Reasoning	.28***
Academic	.38***	Word Knowledge	.34***
Verbal	.36***	Math Knowledge	.25***
Quantitative	.29***	Rotated Blocks	.21***
		Hidden Figures	.21***
		Table Reading	.19**
		Instrument Comprehension	.25***
		Block Counting	.19**
		Aviation Information	.28***

Note. Uncorrected correlations with Final School Grade in 17D Undergraduate Cyber Training Phase II (Manley, 2016). ***p<.001, ** p < .01

Furthermore, there is evidence in a briefing from AFPC for the predictive validity of the AFOQT (Form T) for cyber positions (Air Force Personnel Center, 2017c). Results from a sample of 111 officers in 17D positions revealed that composites of the assessment were predictive of undergraduate cyber training grades (final school grades for Phase 1 and Phase 2), as well as the candidates being in the top 10% of the cyber Phase 2 training course. The relationship between the Verbal, Quantitative, and Academic Aptitude composites and Phase 2 training scores were .39, .29, and .38, respectively. Significant correlations between the Verbal, Quantitative, and Academic composites and average (Phase 1 and 2) course grades were .50, .35, and .48, respectively. The Verbal and Academic composites were significantly related to being in the top 10% of scores for Phase 2 (r = .28 and .19, respectively). In a sample of eligible USAF Officer Candidates (N = 23,151) the Verbal and Academic composites showed more significant Black-White Differences (d = .73 and .85) than Male-Female differences (d = .16 and .34).

3.1.4. Self-Description Inventory (SDI)

The SDI is a trait-based personality assessment that is administered to officer candidates. The early version of the SDI was a 163-item measure of the Big Five personality domains (Openness to Experience, Conscientiousness, Agreeableness, Extraversion, Emotional Stability) (Darr, 2011).

Darr (2011) conducted a meta-analysis to evaluate the effectiveness of the SDI for predicting performance in 20 independent Canadian and U.S. military samples (N = 34,217). Notably, the

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19

meta-analysis included many different adapted versions of the SDI (Darr, 2011). These adaptations included differences in length, translation into different languages, and the use of different formats (e.g., paper and pencil). They found the SDI (or adapted version) had similar, and in most cases stronger, relationships with job performance to those found in civilian positions with other personality measures of the same traits (see Table 20 for a comparison of the effect sizes). Darr's (2011) meta-analyses concluded that conscientiousness and neuroticism are core attributes necessary for general military success, but that specific occupational areas likely have specific traits or facets of traits that can be identified to maximize the variance accounted for in performance specific to that occupation.

Table 20. Comparison of Personality in Military Context and Civilian Context

Trait	Military Context: Darr (2011)	Civilian Context: Barrick and Mount (1991)
Openness	-.01	-.03
Conscientiousness	.35	.23
Extraversion	.19	.10
Agreeableness	.13	.06
Neuroticism	-.22	-.07

Note. The correlations in the table are the corrected correlations between the personality trait and job performance. The Barrick and Mount (1991) correlations are corrected for range restriction and the Darr (2011) correlations were corrected for unreliability in the job performance measure.

The AFOQT Form S version of the SDI (SDI+) had 220 items measuring the Big Five personality domains as well as Machiavellianism and two Air Force related dimensions (Team Orientation and Service Orientation) (Manley & Weissmuller, 2017). Recently, facet-level scales have been explored for the most recent version of the SDI which appears on AFOQT Form T for the purpose of selection and classification of Air Force officer positions. Table 21 lists the personality domains and respective facets (Manley & Weissmuller, 2017).

Table 21. SDI+ Domains and Facets

Domains	Facets
Agreeableness	Team Player
	Pleasant
	Helpful-Altruistic
	Optimism
	Well-Adjusted
Conscientiousness	Achievement Striving
	Order
	Self-Discipline
	Deliberation
	Unconventional
Neuroticism	Stress under Pressure
	Worry
	Temperamental
	Angry-Hostility
Extroversion	Reserved
	Dominance
	Excitement-Seeking
	High-Intensity Pleasure
	Activity
	Spontaneous
Openness to Experience	Creative Ideas
	Reflective
	Scientific Interest
	Cultured
	Imagination
Machiavellianism	Cynical View
	Interpersonal Tactics
	Envious
	Influence Tactics
	Independent

Item, test, and factor-level analyses for the SDI+ on the AFOQT Form T indicated that psychometric properties for both broad dimension and facet scales are acceptable (Manley & Weissmuller, 2017). Results from factor analyses indicated that 30 facets loaded onto 6 dimensions. Internal consistency (measured using Cronbach's α) ranged from very good to excellent, and were .80 to .93 for facets and .91 to .96 for dimensions (Manley & Weissmuller, 2017). Table 22 indicates facets loaded onto each dimension and their corresponding Alpha.

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19

Table 22. Reliability of SDI+ Domains and Facets

DOMAINS & FACETS	α	DOMAINS & FACETS	α
A - Agreeableness	**0.96**	**C - Conscientiousness**	**0.96**
Team Player	0.85	Achievement-Striving	0.87
Pleasant	0.90	Order	0.91
Helpful-Altruistic	0.89	Self-Discipline	0.88
Optimist	0.88	Deliberation	0.90
Well-Adjusted	0.86	Unconventional (R)	0.85
N - Neuroticism	**0.95**	**O - Openness**	**0.92**
Stress-Under-Pressure	0.86	Creative	0.87
Temperamental	0.88	Reflective	0.83
Worry	0.92	Scientific Interest	0.93
Angry-Hostility	0.91	Cultured	0.91
		Imagination	0.86
E - Extraversion	**0.93**		
Reserved (R)	0.91	**M - Machiavellianism**	**0.91**
Dominance-Leader	0.89	Interpersonal Tactics	0.86
Excitement-Seeking	0.92	Cynical View	0.86
High-Intensity Pleasure	0.90	Envious	0.91
Activity	0.81	Influence Tactics	0.82
Spontaneous-Variety	0.80	Independent	0.84

Notes. From Manley and Weissmuller (2017). (R) = Reverse-scored.

Manley (2016) examined the predictive value of the SDI+ for 17D Final School Grades. In a sample of 17D officers, some SDI+ domains and facets were significantly related to final school grade. The uncorrected correlations found are shown in Table 23.

Table 23. SDI+ Prediction of 17D Training Success

DOMAINS & FACETS	*r*	DOMAINS & FACETS	*r*
A - Agreeableness	**.12***	**C - Conscientiousness**	**.19*****
Team Player	.19***	Achievement-Striving	.21***
Pleasant	.11	Order	.10
Helpful-Altruistic	.07		
Considerate	.08	**O - Openness**	**.05**
Hyper-Competitive	-.12*	Creative	.16**
		Reflective	-.06
N - Neuroticism	**-.21*****	Scientific Interest	.10
Stress-Under-Pressure	-.22***	Cultured	-.10
Temperamental	-.17**		
Worry	-.13*	**M - Machiavellianism**	**-.01**
		Envious	-.06
E - Extraversion	**-.06**	Influence Tactics	.04
Unassertive	-.09	Self-serving	.00
Sociable	-.03		
Dominance	.13*		

Note. Uncorrected correlations of SDI+ domains and facets to 17D Final School Grade (Manley, 2016). ***p<.001, **p<.01, *p<.05.

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19

The AFPC provided evidence of the predictive validity of the SDI+ for officer cyber positions (Air Force Personnel Center, 2017c). Results from a sample of 111 officers in the 17D career field indicated that facets of the SDI+ were predictive of undergraduate cyber training grades for Phase 2 of the training course, an average grade for Phase 1 and 2, as well as the candidates being in the top 10% of Phase 2 training. Of the 6 dimensions, Agreeableness, Conscientiousness, Emotional Stability, and Machiavellianism were significantly related to average grades in training. The following facets of the score were significantly related to all 3 criterion outcomes (grades for Phase 2, average grade, and top 10%): Stress under Pressure (-.32, -.36, -.19, respectively), Self-Discipline (.30, .35, .20, respectively), Team Player (r = .26, .33, .20, respectively), Scientific Interest (.21, .29, .13, respectively), Cynical View of Human Nature (-.21, -.21, -.28, respectively), Achievement Striving (.21, .26, .23, respectively), Well Adjusted (.24, .24, .15, respectively), Worry (-.23, -.23, -.21, respectively), and Envious (-.19, -.20, -.12, respectively). Reserved, Interpersonal Tactics, Temperamental, and Angry Hostility each significantly predicted one of the three outcomes and had non-significant relationships (r < 19) with two of the three outcomes. There were no Black-White differences on a custom SDI+ composite (d = .00) and Male-Female differences were small (d = .34).

In sum, evidence from research on the SDI demonstrated that conscientiousness and neuroticism were important personality domains for predicting success in a military context (Darr, 2011). More recent research on the SDI+ replicated the finding that conscientiousness and neuroticism are important in the context of predicting success in a cyber career field in the Air Force and further found that certain facets of the personality domains (i.e., Team Player, Hyper-Competitive, Stress-Under Pressure, Temperamental, Worry, Dominance, Achievement-Striving, and Creative) were significantly related to success in the cyber career field (Manley, 2016).

3.1.5. Air Force Cyber Test

According to the 2017 AFECD, the Cyber Test is a qualifying assessment for several entry-level cyber-related positions in cases where the minimum qualifying ASVAB score is not met (Air Force Enlisted Classification Directory (AFECD), 2017). For instance, if considering the 3D0X2 and 3D0X3 specialties, an ASVAB G score of 64 or greater is required or a G score of 54 and a Cyber Test score of 60+ is acceptable. The 1N4X1A position also offers an alternative lower score on the ASVAB with a Cyber Test score of 60+.

The Air Force Cyber Test (formerly the Information and Computer Technology Literacy [ICTL] test) was developed to be used as a selection tool for entry-level enlisted cyber-related training specialties (Trippe, Moriarty, Beatty, & Diaz, 2014). The test showed adequate internal consistency on both forms A and B in preliminary testing (alpha = .78, .79) (Russell & Sellman, 2009). The test was highly correlated with other predictor scores, especially the ASVAB subtests. The highest correlation of .73 was with the General Science subtest. The Cyber Test showed a non-significant correlation with 1N4X1 final school grades (r =.15, ns). For other cyber positions, the correlation of the Cyber Test with performance ranged from .14 to .46, with a weighted mean correlation of .21. Subgroup differences were the same as or slightly smaller than those generally observed with ASVAB technical knowledge subtest scores.

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19

The Human Resources Research Organization (HumRRO) conducted studies to develop and pilot test a Cyber Test item bank and built parallel test forms (Paullin, Ingerick, Trippe, & Wasko, 2011; Trippe et al., 2014). Form assembly started with 49 items from the original Cyber Test (ICTL), but also included 118 new items that were screened by subject matter experts (SMEs) from the Air Force, Navy, and Army. These new items were based on knowledge, skills, and abilities (KSAs) that SMEs rated as important for cyber-related jobs. The KSA statements involved four broad areas: Network and Telecommunications, Computer Operations, Security and Compliance, and Software Programming and Web Development. Based on the KSA statements included in Trippe et al.'s (2014) report, it is likely the majority of items are knowledge-based. However, the Cyber Test may also measure aptitudes reflective of Information and Technology Aptitude, which was identified as important in Objective 1.

Trippe et al. (2014) provided psychometric information on the newer version of the Cyber Test. Using both Classical Test Theory and the Item Response Theory 3PL (three parameter logistic) model to evaluate the discriminability, difficulty, and pseudo-guessing, two final forms (Form 4 and Form 5) were assembled from the total item pool. Form 4, consisting of 40 items, had acceptable reliability across all levels of theta (.76-.77). The Form 5 solution, consisting of 30 items, also had acceptability reliability results across all levels of theta (.68-.70). Both forms were most reliable at higher levels of theta; thus, it was more consistent scoring individuals with greater ability. So, the higher one's ASVAB (or ASVAB and Cyber Test composite), the more predictive the new form of the Cyber Test (Forms 4 and 5).

The AFPC explored the predictive validity of the Cyber Test in a sample of 1N4X1A (Digital Network Analyst) airmen using data from 2014 to 2016 (Evaluation of 1N4X1A/B Entry Standards, 2017). The Cyber Test did not significantly predict the training Apprentice Course success rate over the MAGE General composite. However, AFPC did note a small sample size ($N = 112$) and previous studies have found that the Cyber Test can significantly predict success in cyber training (Russell & Sellman, 2009; Trippe & Russell, 2011).

3.1.6. Electronic Data Processing Test (EDPT)

The EDPT is patterned after the IBM Programmer Aptitude Test and was initially created to identify airmen who could be trained to handle new data processing equipment as a result of the increasing automation of Air Force records (Lecznar & Klesch, 1965). The Marine Corps also uses the EDPT for classification into jobs in the computer field (Alley et al., 2007). The EDPT has the reputation of being one of the most difficult of the military entrance assessments (Alley et al., 2007). It has four subtests (each with 30 items); Arithmetic Reasoning, Figure Analogies, Number Series, and Verbal Analogies that are administered during a 90-minute test (see Table 24 for descriptions).

Table 24. EDPT Subtests and Definitions

Subtest	Description
Arithmetic Reasoning (AR)	Test measures how well someone can solve problems in arithmetic. The test requires individuals to understand and organize each problem and then to select a mathematical method or formula to solve each problem.
Figure Analogies (FA)	Test measures how well someone can determine the relationship between objects or things. This ability has been found to predict job-related behaviors such as recognizing patterns in information to form conclusions, extracting meaning out of confusion or ambiguity, thinking clearly about complex situations, and speed in learning new concepts.
Number Series (NS)	Test measures how well someone can figure out the pattern in a series of numbers, and to complete or continue that pattern. This ability has been found to predict job-related behaviors such as recognizing patterns in information to form conclusions and speed in learning new concepts.
Verbal Analogies (VA)	Test measures how well someone can to determine the relationships between objects or things.

Note. Definitions from "Sample EDPT Items.docx" document provided by the Air Force.

Validation studies on the EDPT indicate it is a valid predictor of training performance for certain AFSs, however, there is limited research on the incremental value of EDPT over the ASVAB (Alley et al., 2007). The Air Force examined predictive validity of the EDPT in a sample of 113 1B4X1 trainees (Air Force Personnel Center, 2017d). The validity of the EDPT with final school grades for training was significant (adj. $R^2 = .19$, $r = .44$). The EDPT is not as effective as a stand-alone predictor of training success as the MAGE General composite for 1B4X1 (adj. $R^2 = .26$, $r = .51$). However, the EDPT has incremental validity above MAGE (adj. $R^2 = 0.28$, ΔR^2 of 0.02). The small incremental validity could be due to the high correlations between the EDPT and other aptitude tests administered by the Air Force. Specifically, the EDPT scores were correlated from 0.37 to 0.46 with scores on the MAGE, and the EDPT was positively correlated with all of the ASVAB subtests, with individual subtest correlations with the EDPT ranging from 0.21 to 0.50.

3.1.7. Cyber Aptitude and Talent Assessment (CATA)

The CATA battery was developed by the University of Maryland CASL with the goal being to identify and develop measures that would provide incremental validity in assessing cyber aptitude over general intelligence (Cobb, 2016). Moreover, the CATA battery did not include specific skills, as skills may become obsolete and quickly date measures. Notably, the CATA battery did not include measures of verbal or written ability, as these measures are captured well by the ASVAB.

The CATA assessments were designed to encompass the multidimensional nature of cyber roles. Specifically, Campbell, O'Rourke, and Bunting (2015) proposed a schematic where cyber

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19

positions differ along two dimensions, 1) proactive versus reactive thinking, and 2) real-time versus deliberate action. Proactive thinking requires thinking ahead (anticipating) possible outcomes of actions (e.g., profiling targets and their activities), whereas reactive thinking involves recognizing problems and reacting to them (e.g., analyzing digital forensic evidence). Real-time action involves quick decision-making to resolve issues in a timely manner (e.g., maintaining constant awareness of threats in a highly dynamic operating environment), while deliberate action involves deferring decision-making until all the information is gathered (e.g., collecting and processing systems in order to exploit, locate, and/or track targets of interest).

CATA measures 15 different cyber abilities and traits (see Table 25). These 15 constructs are categorized into five broader sections in the CATA assessment: Critical Thinking, Deliberate Action, Real-Time Action, Proactive Thinking, and Reactive Thinking (see Table 26 for the mapping of the constructs into respective sections). The Critical Thinking component measures working memory and reasoning and is intended to predict across cyber professions. The other four sections correspond to the dimensions outlined by Campbell et al. (2015) and are intended to be job-specific, based on the profile of the cyber position among the dimensions.

Table 25. CATA Traits and Definitions

Aptitude/Trait	Description
Anomaly detection	the ability to detect information that is anomalous in a larger context, such that it does not conform to the expected pattern
Complex problem-solving	the ability to learn and effectively manipulate systems which are complex, opaque, and dynamic
Creativity (convergent thinking)	the ability to explore a variety of solutions by forming connections between concepts that are typically weakly related or unrelated, and ultimately to hone in on one single, correct solution to a problem
Mental model ability	the ability to construct abstract, internal representations of a situation, real or imagined, derived from a narrative or other form of input
Modeling program execution	the ability to scan incoming information, detect patterns and react quickly
Need for closure	the need to arrive at a solution during problem solving
Need for cognition	the degree to which individuals enjoy participating in mentally demanding tasks.
Pattern recognition and scanning	the ability to scan incoming information, detect patterns and react quickly
Psychomotor speed	the ability to respond quickly and to control the speeded motor response in the face of interference.
Resistance to interference	the ability to respond quickly and accurately in the face of proactive interference
Rule induction	the ability to determine the rules that govern a pattern
Spatial visualization	the ability to form and manipulate visuospatial representations
Tolerance for risk	represents the likelihood of an individual to be risk taking or risk averse, and is a factor known to influence decision making
Vigilance	represents the ability to remain vigilant or sustain attention during a task that occurs over a prolonged period of time
Visuospatial working memory	the workspace for briefly holding and manipulating information from the spatial domain

Note. All CATA traits and aptitudes map onto those identified as important for cyber positions in Objective 1.

Table 26. CATA Constructs Mapped onto Sections

	Section				
Construct	**CT**	**DA**	**RTA**	**PT**	**RT**
Visuospatial working memory	X				
Complex problem-solving	X				
Rule induction	X				
Spatial visualization	X				
Need for closure		X			
Need for cognition		X			
Tolerance for risk		X			
Pattern recognition and scanning			X		
Psychomotor speed			X		
Resistance to interference			X		
Creativity (convergent thinking)				X	
Mental model ability				X	
Modeling program execution				X	
Anomaly detection					X
Vigilance					X

Notes. CT = Critical Thinking, DA = Deliberate Action, RTA = Real-Time Action, PT = Proactive Thinking, RT = Reactive Thinking. Based on information from O'Rourke et al. (2017).

Pilot testing of measures in the CATA, O'Rourke et al. (2017) reported internal consistencies ranging from .61 (Matrix Reasoning measure of Rule Induction) to .90 (Dynamic Systems Control measure of Complex Problem Solving). There has been validation work on the CATA, though the results should be interpreted with caution as the Manpower Accession Policy Working Group (MAPGA) reviewed the CASL CATA validation study and found several flaws that affect the interpretation of results. The USAF is conducting additional research to examine the psychometric properties of a few of the CASL CATA subtests.[2]

3.1.8. Air Force Multitasking Test

The Air Force Multitasking Test (MTT) uses a program called SynWin (Elsmore, 1994), which has four windows to allow for four tasks to be presented simultaneously (Barron & Rose, 2017). The four windows can display any combination of the four tasks; a simple memory task, an arithmetic computation task, a visual monitoring task, and an auditory monitoring task (see Table 27 and Figure 3 for more details). The MTT measures several abilities that are critical for pilot performance, including working memory, number facility, oral comprehension, and visualization (Barron & Rose, 2017).

[2] T. Carretta, personal communication, January 14, 2019

Table 27. Air Force Multitasking Test Tasks

Task	Characteristics	Location in Figure 3
Memory	A list of letters is presented then disappears. After a delay, a letter is presented, and the test-taker determines whether or not the letter was present in the initial list.	Upper right quadrant
Arithmetic Computation	Test-takers sum three-digit numbers	Upper left quadrant
Visual Monitoring	Test-takers monitor a needle on a fuel gauge and click to reset fuel when the needle gets close to empty.	Lower left quadrant
Auditory Monitoring	Test-takers click when they hear their call sign but ignore other call signs.	Lower right quadrant

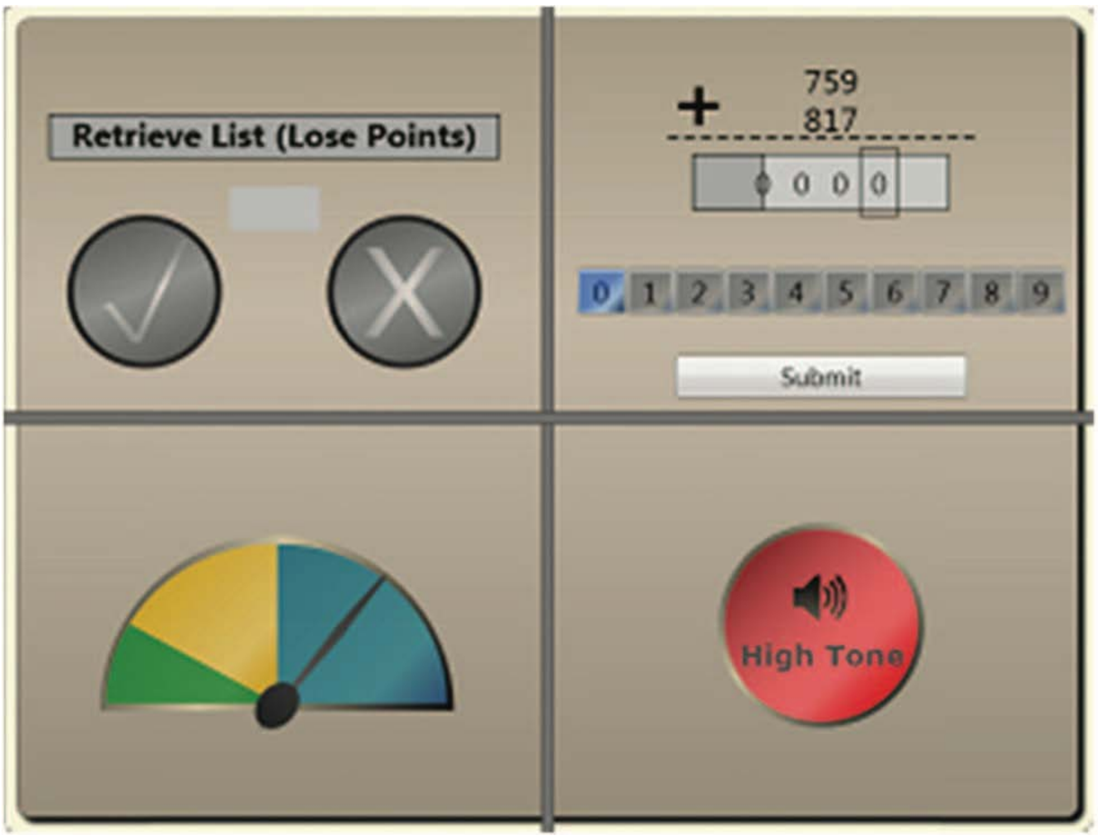


Figure 3. Multitasking Test Screen from Barron and Rose (2017)

The Navy has used the SynWin program to measure multitasking (Hambrick et al., 2011; Poposki, Oswald, & Brou, 2009) as a criterion, while the Air Force has evaluated the MTT as a predictor of performance (Alley et al., 2007; Barron & Rose, 2017) and a training method for UAV operations (Casey et al., 2008). Barron and Rose (2017) found that the post-practice MTT scores of multitasking performance significantly predicted both academic performance (r = .21) and flying performance (r = .23) in a sample of Air Force pilots. Barron and Rose (2017) found adequate reliability for the arithmetic (alpha = .90), visual monitoring (alpha = .72), and

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19

memorization (alpha = .81) tasks. However, the auditory monitoring task demonstrated poor reliability (alpha = .21).

To date, there is no existing research on the effectiveness of the MTT for predicting performance in Air Force cyber positions, though the instrument is promising in that it measures several important aptitudes related to cyber positions. Specifically, Memorization and Working Memory are directly measured in the MTT assessment through the memory task, Number Facility is measured through the arithmetic task, Oral Comprehension is measured through the auditory task, and Time Sharing and Selective Attention could be captured through overall multitasking performance.

3.2 Mapping of DoD Measures onto Cyber Aptitudes and Traits

Next, we determined where gaps exit in current assessment methods. That is, where no current DoD measures mapped onto a particular cyber aptitude or trait that was identified as important in Objective 1. To accomplish this, two advanced graduate students in industrial-organizational psychology sorted the DoD measures into the aptitudes and traits. All discrepancies were discussed to consensus. The above review of DoD measures identified 102 aptitudes and traits measured in the current DoD tests: 9 aptitudes in the ASVAB, 26 personality facets in TAPAS, 10 aptitudes and 6 traits in the AFOQT SJT, 30 facets in the SDI+, 4 aptitudes in the EDPT, 12 aptitudes and 3 traits in the CATA battery, 1 aptitude in the Cyber Test, and 1 aptitude in the Multitasking Test. Table 28 shows the mapping of the important aptitudes and traits onto potential measures. This table includes a “best bet” existing measure as well as one or more alternate measures for each important aptitude and trait, where measures were available. If we did not identify any measures corresponding to a particular aptitude or trait, we labeled that row as a “gap”. Our focus is on the measures reviewed above, which are those currently in use by the USAF.

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19

Table 28. Constructs Measured in Current DoD Assessments

Aptitudes					
ASVAB	**AFOQT**	**EDPT**	**CATA**	**MT**	**CT**
Arithmetic Reasoning	Verbal Analogies	Arithmetic Reasoning	Anomaly detection	Multitasking	CT
Assembling Objects	Arithmetic Reasoning	Figure Analogies	Complex problem-solving		
Automotive & Shop	Reading Comprehension	Number Series	Convergent thinking		
Electronics Information	Word Knowledge	Verbal Analogies	Mental model ability		
General Science	Math Knowledge		Modeling program execution		
	Instrument Comprehension		Pattern recognition and		
Mathematics Knowledge			scanning		
Mechanical Comp.	Block Counting		Psychomotor speed		
Paragraph Comp.	Table Reading		Resistance to interference		
Work Knowledge	Aviation Information		Rule induction		
	Physical Science		Spatial visualization		
			Vigilance		
			Visuospatial working memory		

Traits			
TAPAS	**AFOQT**	**SDI+**	**CATA**
Achievement	Integrity/Professionalism[1]	Team Player	Need for closure
Adjustment	Leadership[1]	Pleasant	Need for cognition
Adventure Seeking	Resource Management[1]	Helpful-Altruistic	Tolerance for risk
Aesthetics	Communication[1]	Optimism	
Attention-Seeking	Innovation[1]	Well-Adjusted	
Consideration	Mentoring[1]	Achievement Striving	
Cooperation		Order	
Courage		Self-Discipline	
Curiosity		Deliberation	
Depth		Unconventional	
Dominance		Stress under Pressure	
Even-Tempered		Worry	
Ingenuity		Temperamental	
Intellectual Efficiency		Angry-Hostility	
Non-Delinquency		Reserved	
Optimism		Dominance	
Order		Excitement-Seeking	

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19

Physical Condition	High-Intensity Pleasure
Responsibility	Activity
Self-Control	Spontaneous
Selflessness/Generosity	Creative Ideas
Situational Awareness	Reflective
Sociability	Scientific Interest
Team Orientation	Cultured
Tolerance	Imagination
Virtue	Cynical View of Human Nature
	Interpersonal Tactics
	Envious
	Influence Tactics
	Independent

Notes. [1]Based on the Situational Judgment portion of the AFOQT test. According to the AFOQT Information Pamphlet (2015), "this composite measures judgment and decision-making in responding to the types of interpersonal situations often encountered by junior USAF officers. Test questions are based on real scenarios experienced by junior officers (O1- O3) requiring core competencies of Integrity and Professionalism, Leadership, Resource Management, Communication, Innovation, and Mentoring. Scored relative to the consensus judgment of identified high-potential USAF officers, test questions were selected for inclusion in the composite based on statistical relationships of scores to cadet outcomes in Basic Officer Training and Field Training."

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19

Table 29 shows the results of the rational sort of the DoD measures listed in Table 22 into the traits and aptitudes identified as important in Objective 1.

Table 29. Best Bet Existing DoD Measures for Critical Cyber Aptitudes & Traits

Type	Construct	Best Bet Existing DoD Measure(s)	Alternate Predictor(s)	Gap?
Ability	Active Learning	TAPAS Intellectual Efficiency	None	No
Ability	Decision Making	None	None	Yes
Ability	Complex problem-solving	CATA (Complex problem-solving)	None	No
Ability	Information Foraging	Cyber Test	None	No
Ability	Problem Sensitivity	None	None	Yes
Ability	Systems Thinking	None	None	Yes
Ability	Fluency of Ideas	SDI+ (Imagination)	None	No
Ability	Category Flexibility	None	None	Yes
Ability	Originality	SDI+ (Creative Ideas), TAPAS (Ingenuity)	CATA (Convergent creative thinking), SDI+ (Unconventional)	No
Ability	Convergent Creative Thinking	CATA (Convergent creative thinking)	None	No
Ability	Mental Model Ability	CATA (Mental model ability)	None	No
Ability	Deductive Reasoning	None	None	Yes
Ability	Inductive Reasoning	CATA (Rule induction)	None	No
Ability	Information Ordering	None	None	Yes
Ability	Mental Agility	None	None	Yes
Ability	Flexibility of closure	CATA (Resistance to interference)	None	No
Ability	Psychomotor Speed	CATA (Psychomotor speed)	None	No
Ability	Anomaly detection	CATA (Anomaly detection)	None	No
Ability	Resistance to Interference	CATA (Resistance to interference)	None	No
Ability	Pattern recognition and scanning	CATA (Pattern recognition and scanning)	EDPT (Number Series), CATA Rule induction, EDPT (Figure analogies), AFOQT (Hidden Figures), ASVAB (Assembling Objects)	No
Ability	Perceptual Speed	EDPT (Figure Analogies), EDPT (Number Series), EDPT Verbal Analogies	AFOQT (Table Reading)	No
Ability	Modeling program execution	CATA (Modeling program execution)		No
Ability	Speed of Closure	None	None	No
Ability	Visuospatial working memory	CATA (Visuospatial working memory)	Air Force Multitasking Test	No
Ability	Spatial Visualization	CATA (Spatial visualization)	Air Force Multitasking, AFOQT (Block Counting), ASVAB (Assembling Objects)	No
Ability	Number Facility	Air Force Multitasking Test, EDPT (Arithmetic Reasoning), AFOQT (Arithmetic Reasoning), ASVAB (Arithmetic Reasoning)	ASVAB (Math Knowledge)	No

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19

Ability	Mathematical Reasoning	EDPT (Arithmetic Reasoning), AFOQT (Arithmetic Reasoning), ASVAB (Arithmetic Reasoning)	AFOQT (Math Knowledge)	No
Ability	Selective Attention	Air Force Multitasking Test	None	No
Ability	Time Sharing	Air Force Multitasking Test	None	No
Ability	Time management	Air Force Multitasking Test	None	No
Ability	Memorization	Air Force Multitasking Test	None	No
Ability	Vigilance	CATA (Vigilance)	None	No
Ability	Oral Comprehension	Air Force Multitasking Test	None	No
Ability	Oral Expression	None	None	Yes
Ability	Written Comprehension	AFOQT (Word Knowledge), ASVAB (Paragraph Comprehension)	AFOQT (Reading comprehension), ASVAB (Word Knowledge)	No
Ability	Written Expression	None	None	Yes
Ability	Information and Technology Aptitude	Cyber Test	None	No
Ability	Instructing	None	None	Yes
Ability	Emotional Intelligence	None	None	Yes
Trait	Achievement	TAPAS (Achievement), SDI+ (Achievement Striving)	None	No
Trait	Adaptability	TAPAS (Tolerance)	None	No
Trait	Analytical Thinking	None	None	Yes
Trait	Attention to Detail	TAPAS (Order), SDI+ (Order)	None	No
Trait	Concern for Others	TAPAS (Consideration)	None	No
Trait	Cooperation	TAPAS (Cooperation), SDI+ (Helpful-Altruistic)	SDI+ (Pleasant)	No
Trait	Dependability	TAPAS (Responsibility)	None	No
Trait	Independence	SDI+ (Independent)	None	No
Trait	Initiative	TAPAS (Courage)	TAPAS (Achievement)	No
Trait	Innovation	TAPAS (Ingenuity), SDI+ (Creative ideas)	None	No
Trait	Integrity	TAPAS (Virtue)	TAPAS Non-Delinquency	No
Trait	Leadership	TAPAS (Dominance), SDI+ (Influence Tactics), SDI+ (Dominance)	None	No
Trait	Need for Cognition	CATA (Need for Cognition)	TAPAS (Curiosity)	No
Trait	Need for Cognitive Closure	CATA (Need for Cognitive Closure)	None	No
Trait	Persistence	None	None	Yes
Trait	Resilience	None	None	Yes
Trait	Self-Control	TAPAS (Even-Tempered), SDI+ (Well-adjusted)	SDI+ (Worry), SDI+ (Temperamental), SDI+ (Angry-Hostility)	No
Trait	Self-Discipline	SDI+ (Self-discipline)	TAPAS (Responsibility)	No
Trait	Situational Awareness	TAPAS (Situational Awareness)	TAPAS (Adjustment)	No
Trait	Skepticism	None	None	Yes
Trait	Social Orientation	TAPAS (Sociability)	SDI+ (Reserved), SDI+ (Independent)	No
Trait	Stress Tolerance	TAPAS (Even-Tempered), SDI+ (Stress under pressure)	None	No

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19

Trait	Team Player	TAPAS (Team Orientation), SDI+ (Team Player)	SDI+ (Reserved), SDI+ (Independent), TAPAS (Cooperation)	No
Trait	Tolerance for Risk	CATA (Tolerance for risk)	SDI+ Excitement-Seeking, TAPAS Adventure Seeking	No

Note. A gap is indicated (Gap? = "Yes") when there are no existing DoD measures for the aptitude or trait.

Most of the important traits identified are currently covered by an existing DoD measure. However, there are clear gaps for Analytical Thinking, Persistence, Resilience, and Skepticism. In addition to some clear gaps where no related DoD measure was found, there are several instances where the coverage is questionable. For instance, Integrity is only measured by TAPAS facets, which is a measure that is currently only used for enlisted airmen. Thus, this measure would not necessarily be available to use as part of a selection assessment for the officer cyber positions. Moreover, there are several instances where the potentially relevant measure is the CATA. However, this measure is still undergoing validation, so the feasibility of using it in a selection context is unknown. Table 30 lists these aptitudes and traits where there are potential gaps due to questionable coverage with DoD measures. Alternative potential measures for these key aptitudes and traits will be reviewed in the subsequent section.

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19

Table 30. Key Gaps: Aptitudes and Traits without Coverage by DoD Measure

Type	Construct	Best Bet Existing DoD Measure(s)	Alternate Predictor(s)
Ability	Active Learning	TAPAS Intellectual Efficiency	None
Ability	Decision Making	None	None
Ability	Complex problem-solving	CATA (Complex problem-solving)	None
Ability	Information Foraging	Cyber Test	None
Ability	Problem Sensitivity	None	None
Ability	Systems Thinking	None	None
Ability	Fluency of Ideas	SDI+ (Imagination)	None
Ability	Category Flexibility	None	None
Ability	Convergent Creative Thinking	CATA (Convergent creative thinking)	None
Ability	Mental Model Ability	CATA (Mental model ability)	None
Ability	Deductive Reasoning	None	None
Ability	Inductive Reasoning	CATA (Rule induction)	None
Ability	Information Ordering	None	None
Ability	Mental Agility	None	None
Ability	Flexibility of closure	CATA (Resistance to interference)	None
Ability	Psychomotor Speed	CATA (Psychomotor speed)	None
Ability	Anomaly detection	CATA (Anomaly detection)	None
Ability	Resistance to Interference	CATA (Resistance to interference)	None
Ability	Speed of Closure	None	None
Ability	Visuospatial working memory	CATA (Visuospatial working memory), Air Force Multitasking Test	None
Ability	Vigilance	CATA (Vigilance)	None
Ability	Oral Comprehension	Air Force Multitasking Test	None
Ability	Oral Expression	None	None
Ability	Written Expression	None	None
Ability	Instructing	None	None
Ability	Emotional Intelligence	None	None
Trait	Adaptability	TAPAS (Tolerance)	None
Trait	Analytical Thinking	None	None
Trait	Concern for Others	TAPAS (Consideration)	None
Trait	Dependability	TAPAS (Responsibility)	None
Trait	Independence	SDI+ (Independent)	None
Trait	Initiative	TAPAS (Courage)	TAPAS (Achievement)
Trait	Integrity	TAPAS (Virtue)	TAPAS Non-Delinquency
Trait	Need for Cognition	CATA (Need for Cognition)	TAPAS (Curiosity)
Trait	Need for Cognitive Closure	CATA (Need for Cognitive Closure)	None
Trait	Persistence	None	None
Trait	Resilience	None	None
Trait	Situational Awareness	TAPAS (Situational Awareness)	TAPAS (Adjustment)
Trait	Skepticism	None	None

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19

3.3 Potential Alternative Measures of Assessment Where Gaps Exist

To identify potential alternative measures where gaps exist, we first conducted a literature review of the related published research and federal and DoD reports. Measures of the important aptitudes and traits found in published research and technical reports are summarized below with corresponding psychometric evidence, where available. In addition to reviewing currently available reports and publications, we contacted I-O alumni working in federal and private sectors. Since cyber security is a relatively new and evolving field, we wanted to get a sense of measures that existed outside of the published literature as well. Twenty-five of the 58 alumni contacted responded (43% response rate) to our inquiry about assessments used for cyber positions. Out of those that responded, only four provided any information on cyber assessments for their organization of employment. The relevant assessments found through our review of the literature and our surveying of the field are summarized below.

3.3.1. Aptitudes

Twenty-six of the 39 cyber aptitudes were identified as potential gaps. Potential alternative measures for these cyber aptitudes are reviewed in the sections below.

Active Learning**.** This construct is defined as, "the ability to understand the implications of new information for both current and future problem-solving and decision-making" ("O* NET Skills", 2018). Though the TAPAS Intellectual Efficiency (IE) dimension conceptually overlaps with the SME sourced construct of Active Learning, IE did not significantly predict JCAC course completion in a sample of 1N4X1A airmen. Thus, the construct may warrant more attention for finding a suitable measure. Raymond James has found Learning Orientation to be a strong predictor of most of their tech roles[3]. The Learning Orientation measure used by Raymond James reflects a tendency to learn through a variety of methods and channels, think abstractly, stay current in the field of study, and engage in self-directed learning activities. This description seems consistent with SME input regarding the importance of Active Learning. Moreover, IBM Talen Management Solutions recently released their Commercial Cyber Aptitude Test (CCAT), which also includes Learning Orientation as a behavioral trait of interest for cyber roles.[4] While Raymond James and IBM could not provide specifics regarding their Learning Orientation assessments, there are several self-report measures for the construct (e.g., Sinkula, Baker, & Noordewier, 1997). Active Learning as defined by the SMEs is conceptually related to Openness to Experience, which researchers found is related to predict contextual performance (Chiaburu, Oh, Berry, Li, & Gardner, 2011) and creativity at work (Dilchert, 2008; George & Zhou, 2001)

Reasoning, Critical Thinking, and Decision Making Aptitudes**.** Several of the aptitudes identified as gaps where there are not existing DoD measures involve reasoning, critical thinking, and decision-making. Potential measures for these aptitudes—complex problem solving, decision-making, problem sensitivity, systems thinking, and deductive reasoning, are reviewed below.

---

[3] A. Newberg, personal communication, May 22, 2018

[4] J. Labrador, personal communication, November 1, 2018. Additional information on the CCAT is available here: https://www-01.ibm.com/common/ssi/cgi-bin/ssialias?htmlfid=67014167USEN

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19

Complex Problem Solving. This aptitude is defined as, "the ability to learn and effectively manipulate systems, which are complex, opaque, and dynamic" (O'Rourke et al., 2017). Complex Problem Solving involves knowledge acquisition and knowledge application (Greiff, Stadler, Sonnleitner, Wolff, & Martin, 2015)(Greiff et al., 2015). This makes the measurement of Complex Problem Solving difficult as it often requires the development of complex tasks or scenarios, such as the "Lohausen" task where the test-taker assumes the role of governing a small city (D. Dörner, Kreuzig, Reither, & Stäudel, 1983). The CATA measure of Complex Problem Solving is undergoing validation. Alternative measures of Complex Problem Solving include tasks involving multiple complex systems, such as the MicroDYN (Schweizer, Wüstenberg, & Greiff, 2013).

Information Foraging. This aptitude involves how people seek, gather, and consume information while adapting to the environmental constraints and changes (Pirolli & Card, 1999). The Cyber Test Development Report (Trippe et al., 2014)listed "ability to search on-line and other resources to obtain information that will help solve a problem" as one of the most highly rated (M = 4.50, SD = 1.37) KSAs for entry-level AF cyber positions. It is unclear to what extent this ability is captured in the current version of the Cyber Test and to what extent this particular aptitude predicts performance in cyber positions.

There are other measures of information foraging that incorporate knowing where to look for information and knowing when to give up (i.e., avoid going down a rabbit hole). For example, Teo, John, and Pirolli (2007) outlined a task where examinees were evaluated on their ability to find the shortest and fastest path through a series of links to a page with the target information. The time to "give up" captures the decision to abandon a search path (i.e., when a person perceived the amount of gain from the path was lower than the cost of moving to a new, hopefully better, source of information). Coovert (2011) detailed a Bayesian application to measuring information foraging for today's warfighter.

Decision Making. This aptitude is defined as "the ability to consider the relative costs and benefits of potential actions to choose the most appropriate one" ("O*NET Skills," n.d.). One identified self-report measure for this aptitude is the Adult Decision-Making Competence index (A-DMC; Bruine de Bruin, Parker, & Fischhoff, 2007). The A-DMC assesses several components of decision-making skills: resistance to framing, recognizing social norms, under/overconfidence, applying decision rules, consistency in risk perception, resistance to sunk costs, and path independence. The validation evidence presented for this measure showed higher-scoring individuals reporting fewer negative life events, suggesting greater decision-making ability (Bruine de Bruin et al., 2007). However, the A-DMC components demonstrated poor internal consistency (.54-.75) and test-retest (.28-.77) reliability values, which may be problematic for its use in a selection context.

Problem Sensitivity. This aptitude is defined as "the ability to tell when something is wrong or is likely to go wrong. It does not involve solving the problem, only recognizing there is a problem" ("O*NET Abilities," n.d.). A closely related concept is situational awareness, defined as "the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future". This

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19

concept has long been considered important in the military context, specifically in terms of pilot performance in Air Force research (Alley et al., 2007; Carretta, Perry, & Ree, 1996). Though there have been recent efforts to assess situational awareness in the context of cybersecurity (e.g., Lif, Granåsen, & Sommestad, 2017; Malviya, Fink, Sego, & Endicott-Popovsky, 2011), there are no widely used measures of situational awareness for cybersecurity with strong evidence of their validity.

Systems Thinking. This aptitude is defined as "the ability to understand how multiple parts of a system interact and influence each other" (Aronson, 1996; Jose et al., 2016). Two self-report measures were identified that claim to assess systems thinking. However, a major limitation with both measures is that they assess systems thinking in a particular, non-cyber related context. The Systems Thinking Scale Revised (STSR; Davis & Stroink, 2016) assess systems thinking in an environmental context. The Systems Thinking Scale (STS; Moore, Dolansky, Singh, Palmieri, & Alemi, 2010) assesses systems thinking in the context of nursing. Based on our survey of the field, there are no currently validated measures for systems thinking as related to systems used in cyber security occupations.

Deductive and Inductive Reasoning. Deductive reasoning is defined as "the ability to apply general rules to specific problems to produce answers that make sense" and inductive reasoning is defined as, "the ability to combine pieces of information to form general rules or conclusions (includes finding a relationship among seemingly unrelated events)" ("O*NET Abilities", n.d.). One widely used measure of deductive and inductive reasoning is included in the California Critical Thinking Skills Test (CCTST; Facione, 1990; Facione, Facione, Blohm, & Giancarlo, 2002), which includes subscales assessing analysis, evaluation, inference, deduction, induction, and overall reasoning skills. The Department of Homeland Security (DHS) uses Simpson and Nester's (2007) Taxonomy for Logic-Based Measurement as a blueprint to develop items that measure deductive and inductive reasoning. The taxonomy provides possible correct and incorrect logical reasoning following a given premise or premises.

Information Ordering**.** This aptitude is defined as "the ability to arrange things or actions in a certain order or pattern according to a specific rule or set of rules (e.g., patterns of numbers, letters, words, pictures, mathematical operations)" ("O*NET Abilities", n.d.). This aptitude can be similarly assessed like category flexibility using the Wisconsin Card Sorting Task (WCST; Grant & Berg, 1948). Cards are assigned to be sorted according to a set of rules.

Mental Agility. This aptitude is defined as "the ability to restructure one's knowledge as an adaptive response to changing situational demands" (Jose et al., 2016). Mental agility has also been described similarly under a different term: cognitive flexibility (e.g., Spiro & Jehng, 1990). One option to assess this aptitude is the Cognitive Flexibility Scale by Martin and Rubin (1995), designed to assess an individual's "(a) awareness that in any given situation there are options and alternatives available, (b) willingness to be flexible and adapt to the situation, and (c) self-efficacy in being flexible" (p. 623).

Idea Generation and Proactive Thinking Aptitudes**.** Most of the aptitudes related to idea generation and proactive thinking overlap with existing DoD assessments. For instance, originality, "the ability to come up with unusual or clever ideas about a given topic or situation, or to develop creative ways to solve a problem" ("O*NET Abilities", n.d.), conceptually overlaps with the CATA measure of convergent creative thinking, the SDI+ measure of Creative Ideas, and TAPAS Ingenuity. The CATA measures of Mental Model Ability and Convergent Creative Thinking are currently being validated.

Fluency of Ideas. This aptitude is defined as "the ability to come up with a number of ideas about a topic (the number of ideas is important, not their quality, correctness, or creativity)" ("O*NET Abilities", n.d.). Carroll (1993) identified idea production as one of the eight second-stratum factors underlying general cognitive ability (National Research Council, 2015). According to Hocevar (1981), fluency of ideas (i.e., coming up with "lots of ideas") is a component of creativity. The National Research Council (2015), also identified frequency/fluency and quality/usefulness as the two most important components of measuring creativity. The U.S. Army Research Institute used the Consequences test, which required test-takers to list as many outcomes (consequences) as they could think of to five given scenarios, to measure the fluency of ideas component of divergent thinking (Mumford, Marks, Connelly, Zaccaro, & Johnson, 1998; National Research Council, 2015). They found that divergent thinking provided incremental validity in predicting leader achievement in a military sample above and beyond expertise and general intelligence (measured by the Employee Aptitude Survey) (Mumford et al., 1998).

The Educational Testing Service (ETS) also has examined the predictive validity of idea generation for graduate school success (Bennett & Rock, 1995; Frederiksen & Ward, 1978; National Research Council, 2015). Frederiksen and Ward (1978) found that their idea generation measure, which was used in an experimental section of the Graduate Record Examinations (GRE), was significantly related to multiple measures of graduate student success and was unrelated to GRE scores, suggesting the importance of fluency of ideas for success beyond other cognitive abilities.

Category Flexibility. This aptitude is defined as "the ability to generate or use different sets of rules for combining or grouping things in different ways" ("O*NET Abilities", n.d.). One identified approach to assessing this aptitude is the use of card sorting tasks, one of the most popular being the Wisconsin Card Sorting Task (WCST; Grant & Berg, 1948). This task involves participants being asked to sort a series of cards into one of four groups. However, the rules for sorting the cards into the four groups are not presented. Participants are only told whether a sorting decision was correct or incorrect. Participants then attempt to learn from the feedback the underlying sorting rules. After a number of card sorts, the underlying decision rules change, and participants must learn the new rules based on continuing feedback. Performance on this measure can be assessed by the total of preservation errors (error made from applying old rules) and non-preservation errors during the task. Originally done with actual cards in person, the task itself can be easily administered in an online format.

Perceptual Aptitudes. The perceptual aptitudes (Flexibility of closure, Psychomotor Speed, Anomaly Detection, Resistance to Interference, Speed of Closure, Visuospatial Working Memory, and Vigilance) are covered by the CATA, EDPT, and AFOQT (Table Reading) assessments. The validation work is ongoing for the related CATA measures (Psychomotor Speed, Resistance to Interference, Anomaly Detection, Pattern Recognition and Scanning,

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19

Modeling Program Execution, Visuospatial Working Memory, and Vigilance). The EDPT measures of Number Series, Figure Analogies, and Verbal Analogies appears valid as it predicted final training grade for a sample of 1B4X1 airmen. The AFOQT Table Reading subtest correlated with 17D training scores in a sample of officers. These existing DoD measures should be further vetted for the cyber positions of interest.

Oral Comprehension. This aptitude is defined as "the ability to listen to and understand information and ideas presented through spoken words and sentences" ("O*NET Abilities", n.d.). An early version of the Air Force Multitasking Test measured oral comprehension through test-takers' ability to recognize and react to noise cues (i.e., distinguish between low- and high-pitched tones). This simple task was replaced by a more complex verbal task that requires examinees to monitor spoken "call signs" and acknowledge their call sign when it is presented. While the modified task involves oral comprehension, it is still relatively simple compared with more complex tasks involving the expression of ideas presented through spoken works and sentences. Measures of this aspect of oral comprehension are often included in assessments of intelligence. One widely used measure includes the auditory comprehension subtest of the Kaufman Adolescent and Adult Intelligence Test (KAIT; Kaufman, 1993).

Oral Expression. This aptitude is defined as "the ability to communicate information and ideas in speaking so others will understand" ("O*NET Abilities", n.d.). Speaking ability, in the context of the English language, has been widely assessed with the Test of English as a Foreign Language (TOEFL; Chapelle, 2011). However, this test is designed for individuals for whom English is a foreign language. There is a lack of existing measures of oral expression for individuals that English is their primary language. This is likely due to the difficult and time-intensive nature of capturing and analyzing oral expression. However, advances in natural language processing may help to reduce these costs.

Written Expression. This aptitude is defined as, "the ability to communicate information and ideas in writing so others will understand" ("O*NET Abilities", n.d.). Written expression is traditionally difficult to measure due to the time-intensive nature of scoring written responses. ETS uses automated scoring and natural language processing to implement automated scoring for essays (https://www.ets.org/research/topics/as_nlp). To the extent that technological innovations in methodology can reduce the costs associated with measuring written expression, the addition of essays to selection tests for cyber positions should be considered.

Instructing. This aptitude is defined as, the ability to teach others how to do something ("O*NET Skills", n.d.). Most measures that assess teaching aptitude focus on instructors in the context of education (e.g., Barr, 1940; Torgerson, 1934). However, in the context of the cyber positions, instructing would be used in more of a mentoring capacity. This ability to transfer knowledge and skills to .
others is especially important in the cyber field where technical knowledge and skills are constantly evolving. An aptitude for mentoring or instructing others could be measured through behavioral interviews or situational judgment tests (SJTs) that assess leadership qualities. One alumnus mentioned that at Visa, their cybersecurity organization HR representative uses behavioral interviewing to assess leadership skills for cybersecurity positions[5].

[5] N. Singla, personal communication, May 23, 2018

Emotional Intelligence**.** This aptitude is defined here as the ability to be aware of others' reactions and understand why they react as they do (definition for social perceptiveness, "O*NET Skills", n.d.). There are several self-report emotional intelligence measures available. Two of the most widely used include the Emotional Quotient Inventory (EQ-I; Bar-On, 1997) and the Mayer-Salvoey-Caruso Emotional Intelligence Test (MSCEIT; Mayer, Salovey, Caruso, & Sitarenios, 2003). Both measures are widely used in organizational contexts. The main difference between these measures is the MSCEIT is more abilities-based, where the EQ-I questions are designed with correct/incorrect answers.

3.3.2. Traits

The majority of the important traits identified are covered by personality facets in the SDI+, TAPAS, or both. In cases where the trait is only measured by TAPAS or SDI+ (Adaptability, Concern for Others, Dependability, Independence, Initiative, Integrity, and Situational Awareness), the feasibility of using the SDI+ measure for enlisted or the TAPAS measure for officers would need to be evaluated in order to determine that there is no gap in coverage. The traits that were not covered by any existing DoD measures are reviewed below.

Analytical Thinking**.** This trait is defined as, the degree to which individuals analyze information and use logic to address work-related issues and problems ("O*NET Work Styles", n.d.). Epstein, Pacini, Denes-Raj, and Heier (1996) measured analytical-rational thinking using items from a Need for Cognition scale, with items such as, "I would prefer complex to simple problems." Allinson and Hayes (1996) developed a 38 item Cognitive Style Index designed to measure where people fall on the dimension of analytical versus intuitive thinking.

Persistence and Resilience**.** Persistence is defined as, the degree to which individuals have persistence in the face of obstacles ("O*NET Work Styles", n.d.). Resilience is defined as, "the capacity to rebound or bounce back from adversity, conflict, and failure or even positive events, progress, and increased responsibility" (Luthans, 2002). Both these traits closely relate to the concept of grit, defined as 'perseverance and passion for long-term goals… working strenuously toward challenges, maintaining effort and interest over years despite failure, adversity and plateaus in progress" (Duckworth, Peterson, Matthews, & Kelly, 2007). A scale developed by Duckworth and colleagues (2007) demonstrated predictive validity for success in Ivy League undergraduates and West Point cadets. However, it is worth noting that subsequent research has noted its high correlation with other personality traits, such as conscientiousness, and possible construct redundancy (Credé, Tynan, & Harms, 2017). Thus, it may be pragmatic to use pre-existing measures of conscientiousness, such as ones included in the TAPAS and SDI+, for measuring this particular trait.

Skepticism. This trait is defined as the degree to which individuals distrust others (Turner & Valentine, 2001). The development of a scale for professional skepticism has been researched in auditing literature as a multidimensional individual characteristic that can be both a trait and a state (Hurtt, 2010; Hurtt, Brown-Liburd, Earley, & Krishnamoorthy, 2013). Turner and Valentine (2001) developed an 11-item self-report scale of cynicism for use in moral decision-making. The importance of skepticism for USAF cyber career fields was brought to us by a SME

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19

who spoke about the motto in the intelligence career fields, “trust but verify”. The SME pointed out that this trait is useful to a point but can also become a hindrance if it is taken too far and a person does not trust any information, suggesting there may be a curvilinear relationship with performance. There are no current measures we are aware of that capture skepticism in the intelligence or security context.

3.3.3. Other Measures

Thorough our research, we encountered several gamified assessments that are used for the selection of cybersecurity personnel, such as assessments developed by Hazardous Software, and a job simulation assessment called “CyberSim” that was developed by Booz Allen Hamilton. Since Objective 3 will cover the potential use of serious games to assess important cyber aptitudes and traits, we will cover these gamified assessments in later sections.

This section summarized potential alternative measures and measurement issues associated with cyber aptitudes and traits that are not currently measured by DoD tests. Some of the important cyber aptitudes and traits are not covered by any current DoD assessments. Other cyber aptitudes and traits are covered by current DoD tests, but there is either no psychometric or validity evidence available (e.g., CATA measures) or the current psychometric or validity evidence suggests it may be worth evaluating potential alternative measures. For many of the cyber traits, only the SDI+ (an officer test) or TAPAS (an enlisted test) measure a corresponding trait. This could be an issue for ensuring coverage across both the officer and enlisted AFSCs. Where there are no feasible alternative measures, or where measurement gaps remain, it will be useful to consider how serious games could be used for assessment.

3.4 Summary

Our review of the literature considered current assessment approaches for the aptitudes and traits important for competency in cyber occupations. While the DoD and civilian sectors each provide a myriad of assessments, there still remain assessment gaps for those aptitudes and traits related to competency in cyber AFSCs. Serious games could provide an approach to fill these deficiencies. Serious game technology would provide a face valid environment for assessment while simultaneously enabling the assessment of multiple aptitudes and traits in a relatively short amount of time.

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19

4.0 OBJECTIVE 3 – RECOMMENDATIONS FOR SERIOUS GAMES

We reviewed the literature on serious games related to the aptitudes and traits of interest to cyber occupations. While key word searches yielded thousands of hits, relatively few articles actually reported on an assessment of the aptitudes and traits of interest. Those relevant studies and serious games are described below. Following that, we considered serious games provided by commercial vendors. A similar state of affairs exists in that the commercial description reports a game to be able to assess nearly anything of interest; yet no validity data are provided. We then move to describing how a serious game could be developed to measure the constructs of interest. We identify the top aptitudes and traits we believe would be most beneficial to assess in a serious game. Development assessments of those constructs should follow rigorous psychometric strategies employing confirmatory factor analysis and multi-trait-multi-method modeling to demonstrate construct, convergent, and discriminate validity. Similarly, human factors assessment needs to demonstrate that the game is appropriately developed from that perspective.

4.1 Literature Review

We conducted a review of the literature on serious games related to the constructs of interest. Computer-based literature searches of Web of Science, Defense Technical Information Center (DTIC), Institute of Electrical and Electronics Engineers (IEEE), and Association for Computing Machinery (ACM) were used to locate relevant studies on the constructs of interest. Additionally, journals highly related to serious gaming, such as The International Journal of Serious Games (IJSG) and Simulation & Gaming, were manually reviewed for pertinent studies. To be included in the initial review, each document had to contain terms relevant to serious games or simulations and the assessment of psychological constructs. Specifically, several searches were conducted of each database/journal. These were constrained using two criteria: (1) to identify articles that contained at least one term related to serious games ("serious gam*", "game engines", "business games", "technology-enhanced assessment", "stealth assessment", "game-based assessment") and (2) contain one of the constructs of interest identified in Objective 1 (see Appendix A for a complete list).

Initial searches resulted in 3,947 hits on articles that included the search terms. To identify relevant studies from these hits, additional inclusion criteria were developed. First, the article had to include an actual measure of the construct of interest. This requirement eliminated many articles that included the constructs outside the context of psychological measurement (e.g., "initiative" as a research initiative rather than a measure of initiative). Second, based on the interest in using the game-based assessment of these constructs in a selection context, articles describing games which directly (or automatically) measured the construct of interest (as opposed to having SMEs observe and score the gamer's performance) were included for further review. This decision was made to further refine the relevant articles because it would be costly and unrealistic to expect to use SME input during the assessment of a large number of candidates in a gaming context.

A detailed screening of the hits based on the inclusion criteria reduced the list substantially. The relevant remaining articles are reviewed in detail in the following section. The review of each game includes a description of the game content, the operationalization and scoring of the

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19

construct(s) of interest, psychometric information when available, and specific considerations for the use of the game in a selection context.

4.1.1. Decision-Making, Adaptability, Collaboration, and Leadership

A notable game-based training platform that assesses decision-making, communication and leadership skills is the *Adaptive Thinking and Leadership (ATL) Simulation* developed by Sandia National Laboratories and Virtual Heroes (Raybourn, Deagle, Mendini, & Heneghan, 2005). This game is used to train soldiers on critical thinking skills, negotiation, information gathering, and communication skills in simulated operational scenarios (Raybourn et al., 2005). Real-time assessments are collected from fellow trainees in the role of "observer" throughout the duration of the simulation (Raybourn, Roberts, Diller, & Dubow, 2008). The aggregated, time-stamped evaluations are displayed after the conclusion of the simulation for a large group discussion that takes place in an after-action review in order to provide feedback to the trainee on what they could/should have done differently (Raybourn et al., 2008). The gaming platform has been used in training at Ft. Bragg for Special Forces soldiers (Williams, Bates, McGovern, Heuring, & Walton, 2006). The utility of this game for the assessment of leadership, communication and/or decision-making skills for selection would depend on the ability to automate the evaluations of performance in the game rather than having several people in the "observer" role evaluate performance.

Lewis, Ellis, and Kellogg (2010) utilized a virtual world, *Second Life*, to measure leadership. They modified an existing puzzle within the virtual world called *Crossing the Ravine*, the goal of which is to collaboratively build a bridge over a ravine. In the original puzzle, colored pieces were assembled into the bridge by having players sit in color-coded chairs that activate corresponding pieces of the bridge). In the modified version, there were two roles to be played: worker and leader. While workers still sat in the colored chairs in order to move the coordinating puzzle pieces, the leader sat in an elevated chair above the workers and instructed them to correctly align the pieces via text chat in the world. Lewis and colleagues (2010) coded the transcripts for each game, establishing how many times the leader communicated, and the communication category for each communication. The three categories were "instructions", "praise", and "other". They also determined the style in which the leader communicated, either imperative or polite. The authors found that previous real-world leadership experience was related to leadership performance in the game (i.e. how well the leaders gave direction).

Swezey and colleagues (Swezey, Streufert, Unger, & Van Rijn, 1985) described the development of a computer-based simulation prototype for the assessment of the complex decision-making skills required of senior Army leaders. Using the *Managerial Assessment and Training Simulation System (MATSS)*, the simulation collects real-time data on participants' actions. Fourteen different metrics for decision-making are collected during the simulation of a political military situation termed the "Yugoslav Dilemma", ultimately producing a "decision-making profile" for each participant. Throughout the simulation, participants stop the game to enter decisions by hitting the "D" key on their keyboard. This information is used to create 14 measures of decision-making. They are: 1) number of decisions, 2) number of respondent decisions and percent of total decisions, 3) number of decision categories used, 4) number of forward integrations, 5) multiplexity F, 6) weight factor in minutes of simulation time, 7) number

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19

of backward integrations, 8) number of unintegrated respondent decisions, 9) quality of integration strategies (QIS), 10) weighted QIS, 11) average response speed in seconds of simulation time, 12) number of serial decision connections, 13) number of integrations planned but not executed, and 14) number of general unintegrated decisions. The content validity was established for this simulation by SMEs' evaluation of the realism of the scenario and the opportunities to demonstrate complex decision-making. The researchers mentioned future research plans for establishing the construct and criterion-related validity of the simulation (Swezey et al., 1985).

Another serious game, *D-CITE* (Decisions based on Collaborative Interactions in TEams), has been used to analyze complex, collaborative decision-making in the context of air traffic management (Freese & Drees, 2016; Schier, Freese, & Muhlhausen, 2016). The game involves working with key stakeholders (the airlines, airport, air traffic controller, ground-handler) from a control center to analyze and make decisions regarding critical events occurring at an airport. Points are maximized by creating response plans that result in the most amount of money and passenger satisfaction points. Overall success is determined by the player's score. There is no direct assessment of decision-making or collaboration.

The Army Research Institute (ARI) described a game-based, computer-mediated platform for assessing leadership adaptability through the *Decisive Action* program (Linkov, Fenton, Satterstrom, Gaskins, & Lewis, 2007). The game places soldiers in the position of a commanding officer of a Joint Force operation during a crisis situation. Throughout the course of the game, there are three "adaptive events" that occur to give the participant the opportunity to express adaptive or non-adaptive behavior. The three events include 1) a minefield that they cannot identify until stumbling upon it, 2) sudden artillery fire, and 3) a ground unit ambush. Adaptive responses to these events would ensure the safety of the soldiers under the participant's command. Once the game is completed, SMEs use after action and situation reports available through the game to grade participants on leadership adaptability. The authors note there are plans to automate the assessment of leadership adaptability through the game engine with metrics such as scores for accomplishing individual objectives, time spent on objectives, and errors in judgment (Linkov et al., 2007).

Bououd and Boughzala (2013) designed a collaborative game in which three IT project managers must share scarce resources (budget and staff members) in order to successfully complete five projects on time. Throughout the game, players must negotiate with other managers to lend or borrow resources. Points are awarded based on collaboration with other project managers, timeliness, quality of the product, relationship with staff members, and exploitation of resources. There are three constructs of interest that could be measured through this game: team player, time management, and decision-making. The team player trait is captured in this game as the number of times a player has given assistance with or without negotiation to other project managers. Time management is operationalized as the number of projects delivered according to schedule and is computed from the number of tasks and projects that are completed on time. Lastly, decision-making, although not explicitly operationalized as such by the authors, can be inferred from player's choices. If gameplay conditions are standardized by fixing the choices of the other two managers, comparable estimates of decision-making ability for individual players can be derived. An optimal series of decisions can be outlined for specific projects (each with its own set of required skills, time constraints, and budgetary demands). Thus, the game can be used to estimate how much a player's choices deviate from the optimal series of decision.

Mohan et al. (2014) developed a serious game to assess the decision-making of physicians who must judge the severity of injuries and choose how to manage them. Their model of physician triage decisions was conceptualized as a product of judgement (information processing and probability estimation) and choice (selection between available alternatives). Of particular interest in the game was the use of heuristics in physician transfer decisions when physicians are under increased cognitive load. Although heuristics allow for faster decision-making, performance is believed to only improve when dealing with representative rather than non-representative cases of trauma. Players were required to evaluate ten cases in which decisions were represented as selections from a pre-specified list of 250 medications, studies, and procedures. The cases ended when physicians either made a disposition decision (admit, discharge, transfer) or the patient died. The influence of heuristics on decision-making was evaluated by comparing the mean number of transfer decisions for representative and non-representative severely injured patients using chi-square tests. The authors were able to support the game's construct validity by establishing the game's external validity, ability to manipulate one task condition via cognitive load, and participants' responses to those manipulations in terms of their consistency with predictions based on cognitive theory. This study demonstrates that experimental manipulation of task conditions can be used to increase cognitive load and assess decision making in scenarios that are representative of those found in real life. In the realm of cybersecurity, a similar format could be used to assess how well cyber operators respond to potential threats by comparing their set of responses to the recommended course of action identified by SMEs. Analytic thinking can also be assessed by the frequency of correct decisions in non-representative cases wherein heuristics must be overridden by more deliberate and careful judgment in order to make the correct decisions.

Jalali, Michael, and Stuart (2017) developed a simulation game to study the ability of decision-makers (experienced professionals versus an inexperienced control group) in handling potential delays in capability development and uncertainties in predicting cyber incidents when building cybersecurity capabilities. In the game, players decide how to invest in building cybersecurity capabilities for an anonymous company by monitoring the effects of their decisions over the course of five simulated years. Throughout each iteration, players must learn what percentage of their resources to spend on cybersecurity and how to distribute those resources among three capability categories (computer-based information systems, cybersecurity capabilities, and cyber incidents). The goal of the game is to maximize profits by finding an optimal balance between profits and money invested in cyber protection, detection, and response capabilities. In other words, players must be able to recognize there is a benefit to early reductions in total profits in order for high profits to be maintained over the long term. Performance depends on three interrelated aptitudes: adaptability, decision-making, and systems thinking. Since the game assesses these simultaneously, it was difficult to disentangle the individual effects of each aptitude. However, the authors were able to distinguish adaptable players from non-adaptable players by introducing a task change. In level one, cyber incidents in each of the simulations followed a fixed pattern so that there was only a single set of optimal values for the three decision parameters. In contrast, cyber incidents in level two occurred randomly, with new and different sets of optimal values for each simulation. They found that experienced high performers in level one usually did not perform well in level two, and vice versa. The results suggest that experienced players, compared to inexperienced players, were less able to adapt to the new pattern. Decision-making and systems thinking, however, were not distinguished in the players' performance as they are necessarily related in this paradigm (since resource allocation decisions are based upon an understanding of how the three capabilities interact over the long run).

Jensen and colleagues described a game called *Follow Me*, a small unit tactical training game for leaders used by instructors and cadets at the United States Military Academy (Jensen, Presnell, Lunsford, & Cobb, 2014). The game implements an Intelligent Game-Based Evaluation and Review (InGEAR) capability which generates automated and tailored feedback in each scenario-based exercise. In the game, players must decide which weapon systems to employ and how to position them during a tactical operation. An overall performance assessment is given for each of the measures in an after-action review (AAR) summary report (mission status, target units destroyed, protected units, ammunition status, suppression status, and fatigue status) based on how often the relevant problem conditions occurred in the exercise. Although the decision and evaluation logic underlying this game is specific to the tactical deployment of weapons, analogous scenarios that better represent the decisions that cyber operators typically face may be created to assess specific aspects of knowledge-based decision-making. Overall, the format and method of assessment presents an efficient method with potential for deriving performance- and behavior-based measures of decision-making across various types of scenarios.

Linehan and colleagues developed a serious game called *DREAD-ED* to train and evaluate group or individual decision making (Linehan, Lawson, Doughty, & Kirman, 2009). Teams must coordinate responses to emergencies (e.g. floods, fires, volcanic eruptions, chemical spills) by successfully communicating unique information, sharing personnel resources, and appraising the many courses of action available before taking action. The goal is to maximize scores on four dimensions of performance in the game (casualties, hazard risk, operations, and public relations rating) while responding to several events that alter the game state in an unpredictable fashion. There are four unique roles within the game (mayor, site manager, emergency services, and safety officer), each with its own special abilities and class of personnel that must be appropriately assembled and deployed in response to each emergency. Decisions are assessed based on final scores and the number of teams deployed during the last round in which such actions are unnecessary or dangerous. Since this game is a prototype, the validity and reliability of this assessment has yet to be established. For use in assessing cyber operators, the task will need to be made more specific to the domain.

Although some articles found in the search did not directly measure relevant traits, they were able to offer important insights hinting at possible modifications for doing so. For example, Gallagher and Prestwich (2012) proposed several game design features that could be used to foster adaptability. Based on the premise of the Wisconsin Card Sorting Task, where participants are forced to determine unknown rules for sorting the card deck in order to demonstrate cognitive flexibility, the authors suggest that a game that forces players to play by rules that are not stated explicitly should create a situation for participants to demonstrate their level of cognitive adaptability. Similarly, when game rules change non-explicitly, it provides an opportunity to test players' cognitive flexibility. The authors also propose that dynamic and shifting environments will foster adaptability. Therefore, games designed to measure decision-

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19

making aptitude or that automatically gather relevant data may be modified to measure adaptability using the aforementioned paradigm. Such modifications may be used to tailor games to represent scenarios that cyber operators typically face.

Additionally, some authors discussed game concepts for the assessments of the constructs of interest rather than games that have been developed. This includes work by Alcaniz, Parra, and Giglioli (2018) who argued for the use of virtual reality (VR) as a tool for the assessment of leadership and managerial skills. Specifically, the authors proposed the use of VR for the simulation of dynamic, complex, and realistic scenarios that require behavioral responses in order to elicit leadership competencies (Alcaniz et al., 2018). In this way, stealth assessment is used by inferring leadership competency levels from behavioral indices. To ensure a valid inference from the behavioral indices, the use of the evidence-centered design (ECD; Shute, 2011) is recommended to establish the relationships between the behavioral responses to the scenarios and the leadership competencies. The authors proposed that leadership competencies could be measured through the decision taken in a scenario, response time for making the decision, and eye-tracking (Alcaniz et al., 2018).

4.1.2. Active Learning

Most of the serious games, simulations, and business games reviewed mentioned the term "active learning" as a way to express that the games are pedagogical in nature and foster greater acquisition and understanding of knowledge. That is, active learning is addressed in the context of creating a learning environment. Although the following studies did not directly measure the active learning trait, they are summarized in order to describe game conditions that are needed to induce and identify behavioral indicators that can be used to derive an actual assessment of active learning.

As an example, one game that emphasizes active learning is Navarro, Pradilla, and Madriñan's (2010) *Comcity*. The game engine is designed for radio planning technicians and uses avatars and metaphors to provide a different way to learn about radio propagation models (an empirical mathematical formulation for the characterization of radio wave propagation). The authors suggest that the use of the game-based stories and missions to simplify these complex aspects of radio planning fosters an environment for active learning. Though this game is used in a training scenario and there is no mention of a game-based measure for active learning, the ability for the game to elicit behaviors that demonstrate active learning provides proof of concept. Invented scenarios and missions that have a significant learning curve are needed for the measurement of active learning within the game scenario.

4.1.3. Memorization, Spatial Visualization, Vigilance, and Selective Attention

Tost and colleagues (2014) developed a 3D virtual environment, *SmartAgeing*, in which users complete a series of five daily tasks meant to detect mild cognitive impairments in memory, executive function, attention, memory, and spatial orientation. The game logs users' actions and computes a single score for each task based on accuracy, completion time, and distance traveled. Users must navigate the virtual apartment (made up of a dining room, sitting room, bedroom, kitchen, and bathroom) in order to complete the tasks described below:

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19

Task 1. Users must find objects in the environment that match a reference image provided on the screen. Objects are found in locations where they would normally be located in real life. For example, a bottle of olive oil would be on a shelf near the stove in the kitchen. The task requires memory, spatial orientation and attention.

Task 2. Users are asked to listen to a radio and click a button each time a specific word is heard. After a few minutes, they are required to water the flowers while continuing to confirm that they have heard the word. Watering the flowers requires users to pick up a watering can, turn on the faucet, fill it with water, turn off the faucet, and finally water the actual flowers. The task requires divided attention and executive planning functions.

Task 3. Users must find the telephone number of a given person, memorize the digits, and use the telephone to dial the number. Additionally, they are told to turn off the television at the beginning of the task and must remember to do so after finishing the telephone task. The exercise involves memory and executive functions and selective attention.

Task 4. The user must identify the objects selected in task 1 amongst a collection of objects.

Task 5. The user must re-identify the selected objects from task 1, but without following instructions and without order. This task is considered a long-term memory exercise coupled with spatial orientation and attention.

A study by Bottiroli et al. (2017) comparing *SmartAgeing* to the *Montreal Cognitive Assessment (MoCA)* (Nasreddine et al., 2005) found support for its validity. Participants were classified into three groups (low, medium, high) according to their *MoCA* scores; within each of these groups, two subgroups (above vs. below median score) were created for each of the six *MoCA* cognitive domains (short-term memory, visuospatial abilities, executive functions, attention, language, and orientation to time and place). *SmartAgeing* indices of performance were compared to each of these groups. Their results indicate performance on the *SmartAgeing* tasks changed according to the functioning of cognitive domains comprising the *MoCA*, suggesting the tasks are proxies for global cognitive mechanisms.

Although the *SmartAgeing* game was designed for early detection of cognitive impairments, the tasks can be modified to be more difficult in order to distinguish the performance of adults with normal cognitive functioning. The data collected by the game can be used in a similar fashion to specifically calculate performance-based indices of memorization, spatial visualization, and selective attention.

Bonnechère and colleagues (2016) examined eight mini-games designed to assess age-related cognitive decline. To evaluate whether or not these mini-games were able to capture different levels of cognitive functioning, players' scores on each were compared to scores of an existing validated measure, the Addenbrooke's Cognitive Evaluation (ACE-R), which assesses five cognitive domains: attention and concentration, memory, fluencies, language, visual and spatial abilities. Three of the mini-games were designed to capture relevant aptitudes (i.e., memorization, selective attention, vigilance) and correlations with corresponding ACE-R subscales are listed below:

- Memory Sweep - Players must memorize the position of highlighted tiles and remember positions when tiles are removed; performance in this game is related to attention (r = .75, $p < .01$) and memory (r = .81, $p < .01$).
- Unique - Players must select the object that is the odd one out; performance in this game is related to selective attention (r = .52, $p < .01$).
- Rush Back - Players must memorize a card and indicate if it matches the card that came before it; this game requires sustained attention (i.e. vigilance; r = .50, $p < .01$) and working memory (r = .66, $p < .01$).

Although the mini-games were designed to detect age-related cognitive changes, the tasks can be modified to be more difficult in order to capture variance among normal working adults. Additionally, analogous forms of these games resemble cyber operator tasks and could be used to increase fidelity and face validity. For example, *Rush Back* could be modified so players must watch a simulation of network traffic and keep a specific user name in mind as they watch patterns of activity unfold to identify a possible attack.

4.1.4. Innovation

The literature search found many hits for studies mentioning "innovation", however, none of these measured innovation as an individual's trait.

Perhaps, the study that came closest to measuring innovation in a relevant context was conducted by Duin, Hauge, Thoben, and Bierwolf (2009). Duin et al. (2009) proposed a framework for measuring disruptive idea generation, which the authors frame as innovation. The serious game they propose is titled *TheTakeover*, and it "is going to be a multiplayer online game supporting a distributed team in an ideation process within an innovation process. The game is based on perspective changes introduced by a fictional takeover of one company by another. Disruptive elements support brainstorming to facilitate the development of new breakthrough ideas." *TheTakeover* follows a previous game the authors created titled *refQuest* (Duin & Hauge, 2008). *RefQuest* also is focused on supporting disruptive idea generation in a cooperative fashion. While the authors state what the forthcoming features of *TheTakeover* will include, we could find no follow-up literature on whether the game was completed or whether it measured innovation as the authors intended.

4.1.5. Situational Awareness (SA)

Collaboration Production Management (COLPMAN) involves players making production and delivery decisions within the supply chain of a large-scale steel company (Nonaka, Miki, Odajima, & Mizuyama, 2016). Five players are needed for this game: one headquarters (HQ), one upstream factory, and three downstream factories. The game is divided into two sessions: a decision-making and a simulation session. The decision-making session begins with an order from a customer that varies by type (5 possible), size (5 possible), and distance of delivery. HQ then assigns downstream factories to produce goods and sets a delivery date for the order. Each downstream factory then sets up both its own schedule for production as well as specific orders to the upstream factory for needed materials. The upstream factory player then creates their own schedule for completing orders back to the downstream factories. Once all schedules are fixed, the simulation session of the game begins. The simulation session simulates production based on former player decisions in the previous session. There are also fluctuations embedded within the simulation session (e.g., express orders, defects, and lead times) players must monitor. After the simulation session, players have completed a period. There are four periods in a term and four terms in one trial. All five players must both keep track of costs within the game (e.g., stock costs, delivery costs, set-up change cost, and penalties for late deliveries) and continuously collaborate with the ultimate goal of maximizing sales volumes and minimizing penalties each trial. To measure both individual and team situational awareness, this game uses a function hierarchy. The hierarchy is organized into three abstraction layers. The top layer is the obtained goals for successful game play. The second layer is the subgoals to reach the top layer goals. The bottom layer is the set of operations to realize the subgoals. Using game data, individual's' awareness of specific parts of the hierarchy can be recorded. Using this method, cognition vectors are defined as an individual's situational awareness. Quantifying these vectors over the course of a game gave a measure of shared situational awareness.

4.1.6. Pattern Recognition

Tsarava, Moeller, and Ninaus (2018) employed *Crabs and Turtles* to measure computational thinking, wherein pattern recognition aptitude is a component. *Crabs and Turtles* consists of three different games: *The Treasure Hunt*, *The Race*, and *Patterns*. It is primarily designed for children at the primary school level", and thus the serious game in its current form is likely not suitable for the purposes of selection of cybersecurity professionals. However, the results of the study hint at potential performance-based measures within the game. For example, the game *Patterns*, in which individual players have to match cards according to certain rules as fast as possible, represents a task that "is closely related to pattern recognition processes that are necessary in coding, for instance, when decomposing problems, generalizing solutions, and forming loops" (p. 10). The authors report findings relating their pattern recognition construct to other constructs not of interest here (e.g., negative feelings).

4.1.7. Systems Thinking

Adis, Wisecarver, Raber, Wind, and Canali (2017) operationalized systems thinking as "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. (p. 3)". While Adis et al., (2017) do not use a game-based assessment in their report, they do recommend using game-based assessment for systems thinking as the recommended next step in measuring the construct.

In *Quest Atlantis: Taiga Park* , several groups of stakeholders (e.g., timber company, a fishing tour company, and farmers) are dependent upon the river for their livelihood The player is presented with five missions in order to determine the cause for the decline in the Taiga River fish population: (1) interview different non-player characters (NPCs), (2) collect water samples, (3) use deductive techniques to test hypotheses from the first two missions, (4) travel to the future using a time machine (in-game) to gain a larger perspective on the root causes of the problem, and (5) use systems thinking to test new hypotheses using the information gained through playing the first four missions, while being able to use the time machine to jump even further ahead in time. Players write and submit short essays and create and submit causal loop diagrams describing their answers, however, nothing is actually recorded by the game itself. From a face validity perspective, the task seems to reflect systems thinking. However, the burden placed on the teacher (or in our case, rater) to monitor the test takers progress and grade the assignment makes this assessment impractical to use in a selection context.

Whalen, Berlin, Ekberg, Barletta, and Hammersberg (2018) use *In the Loop* as a qualitative measure of systems thinking, created to facilitate experiential learning about material criticality and circular economy (CE) concepts. "*In the Loop* is a turn-based, serious board game intended to increase players' recognition of the benefits of CE approaches by illustrating the causes of, potential effects from, and possible solutions for addressing material criticality. The game centers on twelve elements from the European Commission's 'critical raw materials' list" (Whalen & Peck, 2014). The participants in the Whalen study were university students in three successive cohort years of the Research Methodology of Production Projects. *In the Loop* required the players to role play as the CEO of a manufacturing company and make decisions that allow for each to reach seven "progress points". Points are awarded by making strategic decisions that are resource-efficient and by producing the manufacturing products in-game. Resources become scarcer as the game progresses requiring adaptation on the part of the player. Once the game is complete, the students are asked to write reflection essays. Evidence of systems thinking is derived from these reflective essays. Systems thinking is not measured as a continuous variable, but rather by the evaluator judging the essay to recognize the "necessity to handle multiple perspectives at once in a decision or expressed recognition of what could be interpreted as system boundaries, feedback loops, cause-effect relations, or system influences" (Whalen et al., 2018, p. 340). 49% (n = 35) of the participants recognized the game transitioning from a simple linear system to a more complex system that was incorporating unpredictability into the scenario in-game, with 12 of the students recognizing this transition early into their play.

4.1.8. Cooperation

Two studies employed serious games to measure or quantify behavioral indicators of cooperation. Although the focus of both of these studies was not directly on measuring cooperation, the games were able to obtain frequencies of cooperative behaviors that could be used to derive an assessment of the trait.

Powers and Kirkpatrick (2013) implemented a serious game called *Take-a-Chance*, which is an in-person variation of the classic Prisoner's Dilemma. The game takes place within a classroom setting to examine change in trust and cooperative behaviors at two time points during the semester. Students could choose option X (cooperative choice), Y (non-cooperative choice), or No-Play to earn points. Although *Take-a-Chance* was developed primarily to measure a change in students' trust in their classmates (as estimates of the likelihood that others would cooperate in the game), it also measured individual cooperation (as the selection of the cooperative choice, X). Choices reflecting compliance with agreed upon or tacitly understood rules for obtaining optimal outcomes for the team represent behavioral indicators of cooperation; the frequency of cooperative choices over several rounds can be used to derive estimates of individual levels of cooperation. While such cooperative behaviors may be discouraged by the non-cooperative choices of their teammates, those who are higher in trait cooperation are more likely to maintain a cooperative attitude and continue complying with the overall team's expectations. Therefore, standardized scenarios within *Take-a-Chance* may be used to generate comparable estimates of

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19

an individual's level of cooperation based on the frequency of cooperative choices.

Vegt, Visch, Vermeeren, and de Ridder (2016) developed a two-player videogame called *Breakout* in which each player (blue vs. yellow) controls a paddle to bounce a ball against a brick wall that is chipped away with each hit. The objective of the game is to completely break down the wall and earn points for each brick that has been destroyed. A competitive or cooperative variant of the game can be induced by altering goal-related and interaction rules. In the competitive version, players must battle to hit the ball with their own paddles in order to earn points solely for themselves. In the cooperative version, players must alternate with each other to hit the ball so that points will be awarded to each player. Real-time game data and screen captures of all game rounds were compared to self-reports to verify the rules defining cooperative and competitive gameplay elicited distinctive behavioral patterns. The authors also identified specific behaviors that were observed in both game variants and comprised nearly all gameplay: help, agree, ignore, and obstruct. Players in the cooperative condition engaged in significantly more helping (moving out of the way to allow their partner to hit the ball) and agreeing (staying on one side of the screen) behaviors than those in the competitive condition. Cooperative behaviors can also be distinguished from other behaviors by the distribution of paddles over the game area; helping (in which players take turns covering the game area) results in a larger spatial distribution than agreeing (in which players divide the game area). Thus, cooperation can be identified using behavioral patterns.

Cooperation is operationalized as a behavioral pattern characterized by helping (i.e. moving out of the way) and agreeing (i.e. staying on one half of the screen). Although the main focus of the study was to understand the effects of goal-driven rules on player behaviors, the authors devised a method for identifying cooperative behaviors which can be used to obtain frequencies of these behavioral indicators. Standardized scenarios in the cooperative version of the game can be used to obtain comparable estimates of individual levels of cooperation.

4.1.9. Persistence

DiCerbo (2014) examined whether or not *Poptropica*, a popular commercial game for children, would be a suitable assessment of persistence. *Poptropica* is a virtual world in which players explore "islands" with various quests that players can choose to complete. The quests each involve 25 or more steps (e.g. navigating to key areas, collecting and combining items, talking to characters in scripted chats, playing arcade-style games head-to-head). Persistence is operationalized as time spent on difficult tasks, with difficulty defined by the percentage of players who completed the quest. Log files for each attempted quest containing time-stamped event data was collected for each player and used to create a measure of persistence based on (1) total time spent on quest-related events and (2) the number of quest events completed. Although game-based assessment using *Poptropica* resulted in good reliability (alpha = .87), its incremental validity was not assessed by using other validated measures of persistence. The content of the game is likely unsuitable for adults (as it would lack face validity), however, DiCerbo's operationalization of persistence within a game can be applied to more realistic and complex scenarios.

Ventura and Shute (2013) developed a PC video game called *Newton's Playground (NP)* which was designed to help middle school students understand qualitative physics. There are 74 problems in *Newton's Playground* that require the player to draw/create four simple machines: inclined plane/ramps, pendulums, levers, and springboards. The simple machines are created by drawing objects on the screen which ‘‘come to life’’ in 2D to guide a green ball to towards the target. Although relationships are examined in 2D, everything obeys the basic rules of physics relating to gravity and Newton's three laws of motion. A measure of persistence is derived from the amount of time spent on each problem, both solved and unsolved. Difficult problems were identified as successfully completed problems that required multiple attempts, or problems that were attempted but not completed. Persistence is operationalized as the amount of time spent on difficult problems, with greater persistence being characterized by longer times. Since persistence is a response to difficulty, two forms of indicators were examined—time spent on unsolved versus solved problems. The authors validated the game-based measure of persistence against the Performance Measure of Persistence task (PMP; Ventura & Shute, 2013). They found the game-based assessment was significantly related to the PMP and was predictive of learning of qualitative physics; supporting implementation of a real-time formative assessment of persistence to be used to dynamically change gameplay.

4.1.10. Self-Control

Fernandez-Aranda and colleagues (2012) developed a single-player video game, *Islands*, which was designed as a therapeutic tool to remediate attitudinal, behavioral and emotional processes of patients with impulse-related disorders (e.g. dysfunctional emotional regulatory processes and disinhibited personality traits) (Fernandez-Aranda et al., 2012). Players travel to various islands in order to complete challenging mini-games that are designed to encourage players to reduce the strength of strong negative emotional reactions. The game integrates data from biosensors (galvanic skin response, oxygen saturation, heart rate (HR) and HR variation, skin temperature, breathing frequency) and emotion recognition feature extraction algorithms (facial gesture and speech-based) to detect a player's triggered emotions while confronted with specific game situations. The game automatically responds to the player's emotional states by modifying aspects of game play difficulty. Higher undesired emotional and/or physiological reactions are discouraged since they result in greater difficulty (e.g. fish are more difficult to catch, more obstacles appear in the mini-games). In contrast, relaxed and self-controlled reactions are positively reinforced by the game, making the situations easier to handle and the end goals easier to reach.

Operationalization of self-control is not strictly quantified and defined using biofeedback data but is inferred from a player's ability to quell their emotional and physiological response to challenging situations (i.e. reduced impulsiveness or less engagement of quick and unplanned responses). In the context of game-based assessment of self-control, biofeedback and emotion recognition algorithms could be used to quantify how strongly a player experiences negative emotion in response to increasing difficulty, while still maintaining composure and keeping emotions in check. For example, an extremely difficult task could be used to induce a state of high emotional arousal. An individual high in self-control would be able to reduce their emotional response, stay focused, and complete a task more quickly compared to an individual low in self-control, despite comparable levels of an initial emotional arousal.

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19

In conclusion, several of the studies found in the literature review did not use serious games to directly measure the traits and aptitudes of interest. However, a handful of games with relevance to selection of cyber operators have been found.

4.2 Serious Games from Other Government Agencies and Commercial Vendors

In addition to the academic literature review, we reviewed several commercial vendors for the use of serious games. We found commercial serious games through online searches, conference presentations, and personal communications. The serious games that measure cyber constructs of interest are reviewed below. Though there are several serious games that are marketed by commercial vendors for use in selection decisions, the details on the psychometric properties of the assessments (e.g., reliability and validity) are not publicly available, so it is difficult to get a useful assessment of the utility of the games.

PlayAssess is a game-based assessment designed by cut-e (https://assessment.aon.com/nc/online-assessment/gamification/playassess/). *PlayAssess* embeds psychometrically validated assessments into a gamified platform and fictional scenario. Under the pretext of rescuing the top floor of a skyscraper from the control of a fictional alien character named “Odd”, the user is taken into an elevator setting and required to complete multiple tests in order to make it to the top of the elevator shaft. The elevator ascends each time the player completes an assessment. After completing all assessments, the user is rewarded by reaching the top and defeating Odd. Under the guise of this scenario, behavioral, personality, and ability (e.g., verbal, numerical, abstract and logical thinking abilities, reaction time, multitasking) assessments are administered. However, it appears these assessments are very loosely embedded in this scenario as the feedback relies merely on the completion of assessments and the user does not interact with the game to change any scenario outcomes and game-based behaviors are not captured or assessed. While this game does not require users to have any gaming experience and it does not lose any of the assessments’ integrity through integration into the game, it does not provide the user the opportunity to interact with the game and immersion is therefore limited.

Cognify is a game-based assessment developed by Revelian that measures each candidate’s mental agility, attention, cognitive speed, spatial aptitude, and numerical processing ability through a series of seven mini-games (https://www.revelian.com/employer/product/cognify/). Revelian also provides a game-based assessment called *Theme Park Hero*, which assesses mental agility, attention, cognitive speed, spatial aptitude, and numerical processing ability during only ten minutes of play. In taking on the role of the park manager, the candidate’s job is to keep the park operating effectively and to plan and construct a new water park attraction. Attention is assessed through candidates being required to pop specific balloons for a special event while ignoring others. The opportunity to measure numerical processing ability is provided during a part of the game where the candidate needs to calculate the values of various token combinations. Revelian is a well-established psychometrics company in the Australian market (Toivola, 2018). Revelian’s website includes general validity and reliability information for *Cognify*, citing Dr. Richard Landers as having conducted a study comparing *Cognify* to traditional cognitive ability tests, (https://www.revelian.com/employer/cognify-recognition-validation/). According to the website, he found that test takers preferred *Cognify* and perceived it to be a fairer test than traditional assessments. Moreover, Dr. Landers found that "*Cognify* assessed *g* as well or better than many standalone cognitive ability tests."

Arctic Shores currently offers four game-based assessments; *Cosmic Cadet* measures cognition, thinking style, interpersonal style, and delivering results within a spaceship context; *Yellow Hook Reef* measures numerical, verbal, and abstract reasoning through a pirate-themed game; and *Pinnacle Valley* and *Skyrise City* assesses personal style, aptitude, cognition, thinking style, interpersonal style, and delivering results through tasks embedded in two different settings (https://www.arcticshores.com/game-based-assessments/). Similar to Revelian, Artic Shores website claims their assessments are research-based and "validated using thousands of participants following the strict standards of the British Psychological Society", though the details of the validation studies are not publicly available. A doctoral dissertation comparing the validity of vocational interests measured though *Skyrise city* compared with traditional measures found mixed results (Wear, 2018).

Two other technology-enhanced assessments were reviewed that have limited information available on their websites. Design Interactive created *CogGauge* to measure short term memory, attention, and spatial ability through multiple different game scenarios (http://designinteractive.net/coggauge/). Johnston, Carpenter, and Hale (2011) evaluated the test-retest reliability of *CogGauge.* Journey offers game-based assessment of the Big Five personality traits through leveraging Watson-based technology (https://journey.hr/). Their website advertises "The most advanced Psychometric Tool with patented Artificial Intelligence" and has an average test time of twelve minutes.

We also encountered many serious games that organizations are using as recruiting tools. These companies use game-based assessment for recruiting in order to give applicants a realistic job preview (RJP). Use of an RJP during recruiting can reduce candidate attrition during the selection process and further reduce turnover once in the organization (Earnest, Allen, & Landis, 2011).

Rapt Media created a "gamified interactive recruitment experience" for Deloitte, which takes the user through a day in the life of a new Deloitte employee (http://www.raptmedia.com/customers/deloitte/). The user is welcomed to the company through a short video introduction then they begin working. They receive emails and instant messages with realistic previews of the type of work and challenges that would come up for a typical employee.

Similarly, Formaposte French Postal Service uses talent assessment software developed by KTM Advance to allow potential new hires to experience "A Day in the Life" of the French postal service. Formaposte cited that more than 25% of their new hires did not complete their probationary period of hire due to behavioral problems or lack of motivation. After the implementation of the recruitment game, the percentage of employees who quit their job dropped from over 25% to 8% (https://www.furstperson.com/blog/gamification-in-recruiting-its-becoming-more-fun-to-apply-for-a-job).

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19

There are also a couple of cyber-specific recruitment games from IBM and the National Security Agency (NSA). IBM's "Gamified Cyber Attraction Tool" is designed to take users through threat identification and resolution scenarios that will give them an RJP and make sure they are really interested before applying for cyber positions.[6] Similar to the games discussed above, the idea is that if the applicants or recruits are given an RJP before the selection process is completed, they may self-select out of the process rather than starting at the job (or training) and then realizing that it is not right for them and quitting. Through the NSA's Crypto Mobile Game, *CryptoChallenge,* users decode various kinds of cryptographs while under time pressure (https://www.nsa.gov/news-features/press-room/Article/1630529/nsa-launches-new-crypto-mobile-game-app/). *CryptoChallenge* was implemented in support of NSA's cyber-security initiatives to recruit the best candidates for their cyber positions.

Since entry-level cybersecurity positions tend to have high turnover rates coupled with applicants who may be new to the field, the use of serious games as recruitment tools which not only garner interest and engagement from potential recruits but provide RJPs, is an important consideration for this field.

Table 31 below provides a summary of the serious games identified through our literature review and online searches.

Table 31. Summary Table of Serious Games of Interest

Game Name	Construct(s) Measured	Source
ATL Simulation	Decision-making, communication, and leadership skills	Raybourn et al. (2005)
Crossing the Ravine	Leadership	Lewis et al. (2010)
MATSS	Decision-making	Swezey et al. (1985)
D-CITE	Complex, collaborative decision-making	Freese and Drees (2016)
Decisive Action	Leadership adaptability	Linkov et al. (2007)
N/A	Collaboration, teamwork, time management	Bououd and Boughzala (2013)
N/A	Decision-making	Mohan et al. (2014)
N/A	Proactive decision-making	Jalali et al (2017)
DREAD-ED	Group and individual decision-making	Linehan et al (2009)
Comcity	Active Learning	Navarro et al. (2010)

[6] J. Labrador, personal communication, November 1, 2018

SmartAgeing	Attention, memory, spatial orientation	Tost et al. (2014)
N/A	Attention, memory	Bonnechère et al. (2016)
TheTakeover & RefQuest	Innovation/idea generation	Duin et al. (2008)
COLPMAN	Situational awareness	Nonaka et al. (2016)
Crabs and Turtles	Computational thinking, pattern recognition	Tsarava et al. (2018)
N/A	Systems-thinking	Adis et al. (2017)
Quest Atlantis: Taiga Park	Systems-thinking	Shute et al. (1995)
In the Loop	Systems-thinking	Whalen et al. (2018)
Take-a-Chance	Trust, cooperation	Powers and Kirkpatrick (2013)
Breakout	Cooperation/Competitiveness	Vegt, et al. (2016)
Poptropica	Persistence	DiCerbo (2014)
Newton's Playground	Persistence	Ventura and Shute (2013)
Island	Self-control	Fernandez-Aranda et al. (2012)
PlayAssess	Verbal, numerical, abstract and logical thinking abilities, reaction-time, multitasking, personality	https://assessment.aon.com/nc/online-assessment/gamification/playassess/
Cognify	Mental agility, attention, cognitive speed, spatial aptitude and numerical processing ability	https://www.revelian.com/employer/product/cognify/
Cosmic Cadet, Yellow Hook Reef, Pinnacle Valley & Skyrise City	Numerical, verbal, abstract and reasoning, personal style, aptitude, cognition, thinking style, interpersonal style, and delivering results	https://www.arcticshores.com/game-based-assessments/
CogGauge	Short term memory, attention, spatial ability	http://designinteractive.net/coggauge/
Journey	Personality, Needs and Values	https://journey.hr/

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19

4.3 Serious Games as an Assessment Tool for Cyber Aptitudes and Traits

The selection of individuals for careers in cyber is of ever-increasing importance. In this section we lay the groundwork for a philosophy of measurement for the assessment of cyber workers within the context of a serious game. We begin with a discussion of serious games and why, if they are to be used for selection, it is essential to create an experience for the player which has similar psychological fidelity (Kozlowski & DeShon, 2004) to the environment in which their work is to take place. This allows for an instantiation in the player of the constructs to be used on the job and hence can be assessed in the serious game. We then describe the aptitudes and traits proposed for measurement in the serious game. These aptitudes and traits resulted from an extended effort, described above, whereby we identified those necessary for competency across a variety of cyber occupations, and were also rated for importance by subject matter experts. Our list focuses on those four aptitudes and five traits that are presently not well assessed by existing DoD measures. Next, we provide a section focused on measurement issues. We argue it is imperative to develop psychometrically sound measures for inclusion in the serious game. This is done through rigorous statistical approaches allowing us to assess construct complexity of each aptitude and trait via factor analysis by examining unidimensional structures and bi-factor models; and when all measures are developed, it is important to assess individual and joint validity via multi-trait multi-method confirmatory factor models. This two-pronged statistical assessment will help ensure the purity of each assessed construct, resulting in a clearer picture of the unique contribution of the validity of each for predicting cyber performance. A second piece of measurement centers on the importance of usability analysis for the serious game. We argue that determining usability is an ongoing process and needs to be reassessed with each incremental addition to the game. Extensive pilot testing should be employed throughout game development to ensure both game play and measurement systems are functioning as intended. Finally, the serious game should be assessed in a concurrent validation study to ascertain the validity of the traits and aptitudes for predicting competency of performance with existing cyber workers.

4.3.1. Serious Games

Serious games are an evolving tool in organizations and are currently providing many services, such as education and training (Ford & Meyer, 2014; see recent work by Long & Mulch, 2017 on serious games in cyber training), to these organizations. Serious games are continually being developed to meet other needs as well. Our goal is to consider attributes of a serious game to be used as part of a technologically enhanced selection process for cyber operators in several Air Force specialties. It is important to note that game attributes are an issue separate from the problem of identifying the aptitudes and traits necessary for an individual's competency in a cyber occupation. Person attributes are considered in a distinct section below, where we present nine separate aptitudes and traits that could be measured by a serious game. First, we consider those features of a serious game that would prove useful for selecting individuals into cyber occupations.

4.3.2. Features Common to Serious Games

To begin, we set the context by considering common features of serious games. Several books and articles discuss gaming in general (e.g., Schell, 2008) and serious games in particular

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19

(Dörner, Göbel, Effelsberg, & Wiemeyer, 2016). Additional resources are available in the literature we cite. Upon examination, several wide-ranging themes run through these sources. For example, as pointed out by Schell, there is no 'unified theory of game design'. Rather, designing a game is more of an art, where the developer considers the purpose of the game (here selection) and what follows are a myriad of decisions about such issues as look and feel, timing, rules, rewards and so forth. A primary goal of the game is to utilize principles from psychology in order to create an environment whereby the experience of the player results in cognitive and emotional states similar to those experienced on the job (Kozlowski & DeShon, 2004), so those aptitudes and traits of interest can be assessed with high psychological fidelity. This is a subtle yet important point and bears repeating. The game must enable the player to have an experience so the aptitudes and traits necessary for successful (competent) performance in a cyber occupation will emerge and can be assessed in a veridical manor. Assessment in a serious game typically occurs via three primary methods: directly assessed by the game while play occurs, by trained observers, or through the player themselves (e.g., responding to questions during or immediately after play; or debriefing after play) (Coovert, Winner, Bennett, & Howard, 2017; Coovert, Winner, & Bennett, 2017). We will not advocate for utilizing trained observers for assessment, however, due to the many limiting factors associated with the approach (e.g., cost, increased personnel requirements).

In our case, a serious game for assessment of aptitudes and traits for cyber operators serves the purpose of allowing an individual to interact with artifacts during gameplay in a "digital sandbox" so the aptitudes and traits of interest are instantiated and become available for assessment. As such, it is important to keep in mind the fact that the game itself is merely a means to an end. It is the experience during game play the designer cares about; the game enables the experience, but it is *not the experience* (Schell, 2008, p. 10). Schell (2008) goes on to point out that games are excellent at being able to generate certain experiences, and that only games seem able to generate such experiences as accomplishment, freedom, responsibility, and so on (p. 12); while serious games pursue more complex goals such as skill acquisition (Dörner et al., 2016, p. 4). This notion of one's experience being the only reality we can know has deep roots in both philosophy (Bhaskar, 2010) and physics (Resnikoff, 2012), and while a full discussion is beyond the scope of the present work, the basic premise should be kept in mind as one considers developing a serious game for cyber selection.

Generally, a game is a problem-solving activity approached with a playful attitude (Schell, 2008, p. 34) and a serious game is any game used for a purpose other than purely entertainment (Coovert, Winner & Bennett, 2017; Dörner et al., 2016). In our case, selection of individuals into cyber occupations is the other purpose. A serious game will have many of the components of games in general plus some specific actions or attributes, and these are referred to as *characterizing goals* (Dörner et al., 2016, p. 3). Games in general will typically have four core elements (Schell, 2008, p. 41): 1. *Mechanics*-procedures and rules of the game. 2. *Story*-the sequence of events that unfold in the game. 3. *Aesthetics*-how the game looks, sounds, smells, tastes, and feels. 4. *Technology*-any materials and interactions that make the game possible, everything from paper and pencils through complex technologies. None of these elements is more important than another and it is how they come together and interact that determines the player's experience.

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19

Another key element to a game is the theme, the context or frame in which the play takes place. For example, Chess, Monopoly, and Dungeons and Dragons each have a very strong theme to reinforce gameplay. Advantages found in digital games is the ability to use different modalities to reinforce the theme; including lighting, sounds, music, videos and so forth. Games developed for entertainment try to identify themes that resonate deeply with individuals (Schell, 2008, p. 53). Love, hero, and the fight against evil are themes that span generations and cultures and are therefore strong themes (J. B. Peterson, 2002). It is also possible for games to have subthemes, such as coming together to achieve a common goal. This is especially useful in situations like ours where we may wish to assess teamwork or provide a context where the player is working toward a greater common goal.

We have mentioned several key elements in the design of a serious game, but none is more important than its characterizing goal. It is essential the purpose of the game be clearly specified (Schell, 2008, p. 60) so the game play can be operationalized. For the assessment of cyber operators, the characterizing goal is along the lines of: Develop a web-based game that is engaging to players while it simultaneously allows the game to directly measure the following four aptitudes (Active learning, Decision making, Deductive reasoning, and Systems thinking) and five traits (Analytical thinking, Adaptability, Dependability, Persistence, and Situational awareness).

4.3.3. Existing Classifications of Serious Games

We can gain an appreciation for the scope of serious game applications through a cursory exploration of their typologies. Several typologies exist for classifying serious games (Wiemeyer & Hardy, 2013; Wiemeyer & Kliem, 2012). These include serious games that are meant to develop and assess:

- Cognitive and perceptual competencies/skills (e.g., problem-solving, perception)
- Emotional and volitional competencies/skills (e.g., stress control, endurance)
- Sensory-motor competencies/skills (e.g., reaction time, eye-hand coordination)
- Personal competencies/skills (e.g., self-efficacy)
- Social competencies/skills (e.g., cooperation, communication)
- Media competencies/skills (e.g., media knowledge)

There is also a serious games classification system by Ludoscience (2018) which is useful as one can enter keywords and a directory of available games is provided with summaries. For completeness, one may also want to examine the directory from the *Serious Games Association* (Göbel, 2016).

4.3.4. Issues of Performance Assessment during Game Play

Measuring performance of individuals throughout play is a critical issue for successful games. This is the case because performance must be monitored so an appropriate level of flow and difficulty can be presented to the player, thus keeping them engaged without producing unintended emotions such as anxiety. In terms of software development, the flow of a game is essential to keep individuals from becoming bored when play levels are too easy or, on the other extreme, presenting play levels that are too difficult. Linear flows of development are outdated

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19

approaches and little used anymore. A game is more likely to be successful if one takes a spiral model (Schell, 2008, p. 83) of software development. Spiral models are essentially approaches employing iterative design and development. They begin with an initial prototype, test it, revise, test, revise and so forth. This process continues until a level of acceptable play and performance measurement achieved.

It is important to acknowledge human performance occurs at a variety of levels: biomechanical, neurological, and social, to provide exemplars. As such, human performance can be modeled from different perspectives. One major distinction is a focus on that which is observable versus that which is not. Behavior can also be thought of as general versus domain specific. Furthermore, behavior in each case can be divided into state and process models versus continuous models of performance. For a concise summary relevant to serious games see Wiemeyer, Kickmeier-Rust, and Steiner (2016, pp. 275-277).

Finally, it is important to remember performance evolves over time throughout the play of the game. One might consider early play to be a ‘warm up’ period for the player with later play allowing for the emergence of the constructs we wish to measure. Of course, we always need to keep in mind that those traits and aptitudes we are interested in assessing lie behind the observable actions of the player in the game (Coovert et al., 2017). This is a well-known fact of any scientist who deals with latent variables. Behavior is directly observable and measurable, but what we are interested in are those constructs that determine the behavior and influence its emergence; and these constructs are not directly measurable but must be inferred. This inference is often made through rigorous statistical methods and theoretical models. See Coovert, Miller, and Bennett (2017) for an example of inferring trust in one’s teammates from play in a serious game. Trust is computed as factor scores in an autoregressive model over time within a latent change score framework and latent growth trajectories.

4.3.5. Game Engines

Game engines provide the functionality of a game. Major tasks for engines include efficient graphic rendering, player input, file system access, sound, and networking. Static game engines are often used for dedicated purposes but are quite limited. Dynamic engines are independent of the game and platform and thus, for most applications including ours, represent the engine type of choice. Many issues influence the choice of a game engine, see Freiknecht, Geiger, Drochtert, Effelsberg, and Dörner (2016, p. 154) for a list of useful questions to ask when evaluating choices. For our intended use, three engines are on the short list; these being Unity, Unreal Engine (Epic), and CryEngine (Crytek). Resources are widely available to support development in each of these engines. We will also closely look at the AngularJS, and JavaScript framework used by Long and Mulch (2017) in their *CyberWar: 2025* game.

4.3.6. Serious Game Summary

We argue a serious game should be used to enhance the selection procedures for cyber AFSCs. The game will have as a characterizing goal to create an experience for the player whereby the aptitudes and traits to be assessed are instantiated and available for measurement. This experience is created through utilizing the concepts discussed above, such as crafting a digital world with strong themes and a set of artifacts deemed essential to make the experience robust and satisfying. Flow of the game should be monitored so the player is neither bored nor overly challenged. Usability analysis throughout development should ensure the game is solid from the perspective of human factors. Performance measures and construct evaluations should be rigorously developed to ensure their validity from a statistical perspective (discussed below). We now move on to consider the aptitudes and traits proposed for assessment.

4.3.7. Aptitudes and Traits

Our work described above identified the aptitudes and traits necessary for cyber occupations. In addition to specifying the aptitudes and traits necessary for effective job performance, we also examined the literature to pinpoint existing measures of the traits and aptitudes so a gap analysis would reveal which would be candidates for assessment in a serious game. After determining all the traits and aptitudes, a survey was developed to solicit input from Air Force subject matter experts (SMEs) regarding the importance of the traits and aptitudes and asked them to identify those five most important to the job. These ratings were weighted by the number of respondents in each AFSC. Results of our analysis are presented in Table 32 and comprise those aptitudes and traits we propose for measurement in the serious game. See Table 30 above for the complete list.

Table 32. Proposed aptitudes and traits to measure in the serious game

Aptitude	Gap in measurement		Trait	Gap in measurement
Active learning	Yes		Analytical thinking	Yes
Decision-making	Yes		Attention to detail	No
Complex problem solving	No		Initiative	No
Information and technology	No		Adaptability	Yes
Deductive reasoning	Yes		Dependability	Yes
Anomaly detection	No		Stress tolerance	No
Written comprehension	No		Persistence	Yes
Time management	No		Integrity	No
Systems thinking	Yes		Situational awareness	Yes

Note. Importance ratings were obtained from SMEs and weighted by the number of raters. Aptitudes and traits with a "Yes" in the Gap in measurement column are proposed to be measured in the serious game as there is currently not an effective DoD measure for it.

For the aptitudes, we believe it would be most beneficial to focus on active learning, decision-making, deductive reasoning, and systems thinking because they are the most important aptitudes that are not currently covered by DoD measures. Though the TAPAS measure for intellectual efficiency is similar to active learning, there is no parallel assessment for Officers and the Intellectual Efficiency measure does not fully capture our understanding of the active learning construct (as described by SMEs in focus groups). Complex problem-solving and anomaly detection are measured by the Cyber Aptitude and Talent Assessment (CATA) battery. Information and technology aptitude is measured by the Cyber Test. Written comprehension is

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19

measured by the AFOQT (Reading Comprehension and Word Knowledge) and ASVAB (Paragraph Comprehension and Word Knowledge). Time management is assessed by the Air Force Multi-Tasking Test.

For the traits, we believe it would be beneficial to focus on analytical thinking, adaptability, dependability, persistence, and situational awareness as they are the most important traits that have the largest gaps in terms of whether or not they are covered by DoD measures. TAPAS includes measures similar to adaptability (TAPAS Tolerance), dependability (TAPAS Responsibility), and situational awareness (TAPAS Situational Awareness), but they do not have an Officer equivalent on the SDI+. As a result, these traits were considered not fully covered by current DoD measures. Attention to detail is measured by the TAPAS and SDI+ Order facet. Initiative is covered by the TAPAS Courage and Achievement items. Stress tolerance is assessed on the TAPAS Even-Tempered and SDI+ Stress Under Pressure scales. Integrity is captured by the TAPAS Virtue and Non-Delinquency items.

4.3.8. Possible Game Scenarios

Given the background provided on the development of a serious game and the list of potential aptitudes to traits to assess, we now provide four ideas for the context of the serious game. The first would be set in a cyber office environment. The four aptitudes and five traits would be assessed as the candidates moved throughout an inbox of activities representing a 'day in the life' of a cyber operator. This context would provide the highest face validity for candidates. Other scenarios include a cargo airline run, a holodeck, and an escape room. Each of these offer creative opportunities for the game and assessment. Examples of what these might look like for the ladder three scenarios is presented in Table 33.

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19

Table 33. Three Possible Scenarios and Measurement Strategies.

Game	Aptitudes			
	Active Learning	Decision Making	Deductive Reasoning	Systems Thinking
Cargo Pilot - You are the pilot of a cargo aircraft located at a central hub. You must make nine flights (can vary--one for each assessment), at each destination a separate assessment occurs. Game type: Quest to deliver all the cargo.	Upon reaching this destination, the test-taker learns that there will be an inspection of the cargo. The test-taker receives instructions on how to inspect loaded cargo, cargo lashed to decks or in storage facilities, and cargo handling devices to determine compliance with health and safety regulations and need for maintenance. The test taker is then presented with information on each item and the necessary facts, then is quizzed on whether it passes the rules for inspection or not.	You need to consider the various routes for optimal distribution. Upon reaching this destination a variety of cargo must be unloaded. Decisions must be made regarding the tradeoffs in the weight, size and destination of cargo and the tools (e.g., hand truck, forklift) available for use. An optimal solution would exist to which the assessed performance can be compared.	Upon arriving at this destination, the player finds out communication with mission control was compromised. In order to reestablish communication, the player must complete a circuit board that has one correct sequence. Through lit indicators, players through trial and error will learn rules to complete sequence and reestablish communication with mission control. Time to complete sequence and number of errors would be recorded to assess deductive reasoning for players.	The systems task involves optimal aircraft performance measured by remaining fuel and time. Inputs/variables that are manipulated: angle of climb (increased angle = increased fuel burn), air density (higher temperature/height above sea level = decreased performance), cargo loading (how much cargo they take), weather/cross-wind component Create model where they can manipulate those inputs and have to submit a final model after four minutes.
Star Trek Holodeck - The game has a holodeck as a central hub. The player selects nine different scenarios where an assessment occurs during	Upon arriving at a new world, the examinee has to learn about characteristics of the inhabitants of this world in order to complete their mission and go on to the	Scenario where there is an adversarial threat and a decision must be made on what to do. There are multiple stakeholders who are giving their input. Input	In this scenario, in order to survive the world, players must complete an ordering sequence. Based on context of the game, players will be given a set of rules/premises	Systems task involves optimal aircraft performance measured by remaining fuel and time. Inputs/variables that are manipulated: angle of climb

each one. Game type: Quest to survive the nine trips to nine worlds presented by the holodeck.	next world. For instance, they have to learn about characteristics of the species inhabiting this world. They will be quizzed about the characteristics before leaving	needs to be ranked on importance. Ultimately, the player decides. The player's decision is compared to SME ratings of what the optimal decision is in the given scenario.	to correctly order and include/exclude other worlds/planets to form the final correct sequence. This could be the first world in the game, and the correct sequence would grant access to subsequent worlds for the players. Time and accuracy of task will be used for scoring.	(increased angle = increased fuel burn), air density (higher temperature/height above sea level = decreased performance), cargo loading (how much cargo they take), weather/cross-wind component Create model where they can manipulate those inputs and they have to submit a final model after four minutes.
Escape Room - progression through a series of nine rooms. An assessment occurs in each room. Game type: Quest to finish the nine escapes. As implemented once the assessment is done the player will be moved to the next room/assessment.	To escape one of the rooms, the test-taker has to crack a code. The different components of the code are revealed through clues from different areas of the room. For instance, could be that someone has let loose a massive bio-terror attack that's effects are transmitted person to person creating a worldwide pandemic. You are in the room where the antidote was created, but the scientist who created it has died. You need to decipher his clues to figure out the chemical formula for the antidote. (e.g., C16H13ClN2O)	In this world a series of survival decisions are presented to determine an individual's: resistance to framing, recognition of social norms, over/under-confidence, applying decision rules, consistency to risk perception, resistance to sunk costs, and path independence. These seven components follow models of adult decision making (see Bruine De Brruin, Parker, & Fischhoff 2007; Tversky & Kahneman 1992; Kahneman, 2011)	In this room, there are clues and rules that the player must use deductive reasoning for. Clues will allow players to obtain keys (5 total). The final task within this room in order to advance to the next room will require using all collected keys and ordering them in the correct sequence to get out the door. There will be rules for how they will figure out the sequence for the final door.	Systems task involves optimal aircraft performance measured by remaining fuel and time. Inputs/variables that are manipulated: angle of climb (increased angle = increased fuel burn), air density (higher temperature/height above sea level = decreased performance), cargo loading (how much cargo they take), weather/cross-wind component Create model where they can manipulate those inputs and they have to submit a final model after four minutes.

Traits				
Analytical Thinking	Adaptability	Dependability	Persistence	Situational Awareness
Changing up the formula of the game a bit for the analytical thinking task. Once the pilot returns to the hub for his/her next assignment, the player encounters an issue with how to get the next assignment because the computer that assigns the next task is hacked and unreliable. The player must solve a logic puzzle to get the computer back on-line to assign them their next destination. The time needed to solve the puzzle will be the objective score on the task, with the option of calling in an NPC to solve the puzzle for them after 5 minutes (in which case, they would receive a score of zero for non-completion of task).	After a few flights, an unforeseen event will require players to change their strategy for determining the order in which cargo must be delivered. For example, the most effective route may need to be determined using a different set of weights and/or inputs (e.g. inclement weather; prioritization of a piece of cargo for emergency delivery; following a required order when delivering to a series of destinations). In every scenario throughout the task, the same types of information will be presented. Before the task change, each piece of information will have optimal and consistent weights. After the task change, players must adapt by altering how much weight is given to information. Adaptability can be assessed according to the (i) amount of performance decrement after the change and (ii) how	Upon arrival at a destination the player starts to unload the cargo and gets an incoming call from another cargo aircraft. The other aircraft has broken down on the other side of the island. They have help on the way but you are closer. The player is then given a choice to either help the other aircraft deliver their cargo or not. They are either presented as both favorable (with gains and losses) or the helpful option as less favorable. The reason for this being, so the player does not think helping the other aircraft is what is expected/wanted from them. (Will also pilot test if a medical emergency would be a reasonable alternative task).	v1: During each of the flights, the test-taker has the ability to earn extra points by helping with supplies to the different locations. The idea would be to allow players the opportunity to complete a difficult task where they can drop-off after completing just one piece of it, but those who are persistent would continue to complete all the tasks. v2: The test-taker completes as much of the first task as they can, but the task is structured such that they cannot finish. Then they are taken through the other tasks and with any time remaining on a completed task, they can go back and complete the first task	The test-taker is responsible for monitoring the ADS-B input for where they are in a congested airspace. We'll blank the screen and have the test-taker reproduce local targets.

	quickly performance is recovered.			
The player is presented with three screens and has to figure out the "language" of the puzzle is to draw a line that is the mirrored image of what is already presented on the screen, while avoiding obstacles in the puzzle environment. Similar to the picture presented below.	Their trips will involve selecting resources to bring on each excursion (e.g. food, fuel, medicine, skilled workers, various tools and equipment), foraging for supplies, fending off aliens, avoiding obstacles; before the unforeseen change, they will be able to use a relatively simple and routine strategy for successfully reaching their destination. For example, they may be no penalty for heavily stocking up on supplies for each excursion, supplies will be in locations indicated by similar cues in the environment, and attacks will be easily identified from afar. After the change, they will need to switch from a passive strategy, to a more responsive and defensive one. Overstocking will now slow their movement; mechanical or equipment failures will force them to infer information about their status by other means; additional supplies will not be consistently indicated by environmental cues; attacks	The player is presented with a list on ten tasks when they get to the world. The player is then asked to choose which tasks they will complete (all tasks seem to have even difficulty/time commitment). They are in the world for a given amount of time (ex. 10 minutes) in which the average number of tasks completed is 5. Dependability is measured by how many tasks are completed in comparison to how many tasks they chose.	On the way to their destination world, they will be required to rescue at least 1 (of several) groups of colonists who are stranded on different planets along the way; the groups vary in the amount of difficult involved in rescuing them. Players will be given additional points and resources for each rescue as incentive to attempt additional rescues. Persistence will be assessed according to the (i) amount of time spent on these difficult rescue missions and (ii) number of rescues completed. The rescues can be determined to be "difficult" based on the proportion of players who were able to successfully complete a rescue mission.	Primary task on each world will be to retrieve an important artifact while avoiding hostile aliens who are converging on their location. The secondary task will require players to rescue another, equally important item or person who must be revived, repaired, excavated, etc. Players must work on the secondary task while keeping in mind where the hostile forces are currently located and how much time is left.

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19

	will be more aggressive, harder to detect, and require different defensive maneuvers. Adaptability can be assessed according to the (i) amount of performance decrement after the change and (ii) how quickly performance is recovered.			
In the escape room, in order to obtain one of the deductive reasoning clues that will give the next clue for obtaining a key, the player is presented with a logic puzzle similar to the puzzle pictured below. Once the player has passed the analytical thinking task puzzle, the clue will then be presented so the deductive reasoning task can be evaluated. Scoring can be conducted by the time it takes to complete the puzzle.	The assessment of Adaptability can be done by changing the nature of the clues so that different decision rules must be used. Some objects may actually be distractors meant to mislead players, requiring them to use multiple sources to verify the information provided. Clues will also be found in different places and will require a different type of interaction. For example, instead of clicking on objects to reveal a description, players will need to combine objects in order for a clue to be presented. Or, one object will reveal half of a clue, and another object must be found to complete that clue.	The player has just advanced to the next room and gets a message from a NPC in the room he was just in. The player is asked to assist the NPC in completing tasks he just did. The player is then presented with a choice to either complete the task he is working on and then respond with help or pause his current tasks and respond immediately. The situation has trade-offs but ultimately does not help or hinder game completion.	In the first room, there will be an impossible puzzle that they have the option to complete for extra points. They will have the option to go back and reattempt with any remaining time.	The test-taker is informed that they drank poison and they only have a certain amount of time to find the antidote (primary task) and they also need to find the key to get out of the room (secondary task). There are two boxes with keys to get out of the room, one is more difficult but guaranteed to get them out of the room. The second is easier, but there's only a 50% chance to get out of the room and if they don't solve it, they have to go to a third box. Manipulate time left (e.g., take away the timer)

4.3.9. Aptitude and Trait Summary

Now that the aptitudes and traits to be measured in the serious game have been identified, we move to discuss issues associated with their development. It is essential to take a rigorous statistical approach so construct valid measures can be employed in the serious game. This is essential, because it does not matter how clean, slick, and useful the game appears, if we are not reliably assessing the constructs, we will not be able to determine the extent of criterion space of cyber operator performance being explained by those constructs and their utility for selection of individuals into cyber occupations would remain unverified.

4.5 Measurement

This section considers the issue of measurement. Our job is to measure separate aptitudes and traits throughout game play. As such, construct definition, operationalization, and measurement each hold a central role in establishing the validity of our measures. Below we describe the statistical process that should be followed employing factor analysis, analyzing for bi-factors, and confirmatory analysis of multi-trait multi-method (MTMM) models. This rigorous methodological approach needs to be taken to ensure we avert a problem often associated with serious games where only one or two factors emerge due to high intercorrelations among measures[7]. Furthermore, the field has recently acknowledged the perils associated with employing the incorrect model, as is the case when a unidimensional construct is used when a bi-factor construct is appropriate (Gonzalez & MacKinnon, 2018; Jennrich & Bentler, 2011; Muthén & Muthén, 2015; Reise, Waller, & Comrey, 2000) and when any structure other than MTMM is employed when MTMM is appropriate (Castro-Schilo, Widaman, & Grimm, 2013).

4.5.1. Development of Measures to Assess Aptitudes and Traits

The role of sound psychometric analysis in this work cannot be overstated. The goal is to create a serious game that imparts an experience in the player such that it is possible to critically assess nine separate constructs (four aptitudes; five traits) emergent during play. This can be done by employing a combination of measures directly gathered by the game (e.g., task accomplishment) and those obtained from player responses. In the case of the latter, imagine pausing the game in order for the player to respond to inquiries to assess the level of player situational awareness.

We will employ factor analysis to establish the dimensionality of the constructs measured in the study. Traditional factor analysis is well understood (see Comrey & Lee, 2013; Coovert & McNelis, 1988), and will not be covered here beyond reviewing the structure of a factor so it can be contrasted with the bi-factor model, discussed below.

[7] T. Carretta, personal communication, November 6, 2018

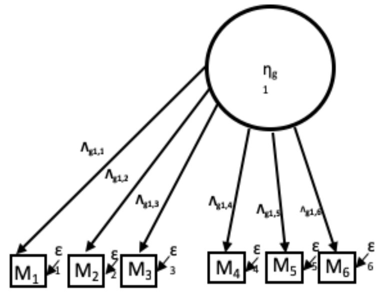


Figure 4. Latent construct with six measured variables

In Figure 4 above uses notation from structural equation modeling where circles are latent constructs, boxes reflect measured variables; η represents a latent construct (the subscript g refers to a general factor and s (in Figure 5 below) refers to one or more specific factors; Λ are factor loadings; and ε refer to errors. Figure 4 may represent a situation where we are measuring the construct *trust-in-technology* with six items from Mcallister's (1995) trust scale.

4.5.2. Bi-Factors

Recently, enthusiasm has developed among methodologists for the assessment of measures via bi-factor modeling. Originally proposed by Holzinger and Swineford (1939) it was reintroduced to the field by Gibbons and Hedeker (1992). It continues to gain wide acceptance as it appears to be a model that explains the structure of many measures, often incorrectly considered unidimensional (Jennrich & Bentler, 2011; Muthén & Muthén, 2015; Reise et al., 2000) including mediational relationships (Gonzalez & MacKinnon, 2018).

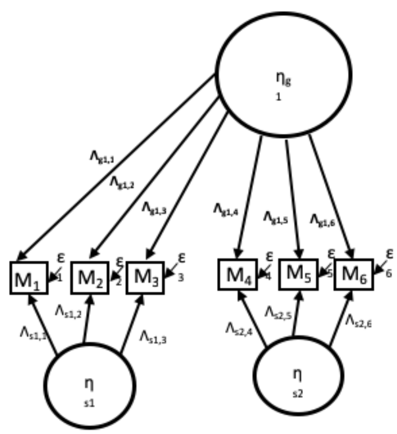


Figure 5. A bi-factor model

The bi-factor representation depicts one general factor and one or more specific factors. It is not a hierarchical structure where one or more specific factors are nested within a general factor. Rather both types of factors (general and specific) are on the same plane and determine the observed score. For our *trust-in-technology* scale, a bi-factor representation (Figure 5) would have one general trust factor, and two specific facets--cognitive trust and affective trust.

Application of the approach is to consider two different models when it comes to analysis, one where the construct is unidimensional and a second where it is functioning as a bi-factor model. Identifying the correct factor structure essential because the general factor and each specific factor (facet) may have different influences on the criterion (Bollen & Lennox, 1991; Gonzalez & MacKinnon, 2018; Reise et al., 2000). Thus, when using them in a prediction context it is essential that the scores not be conflated with one another.

4.5.3. Convergent and Discriminant Validity Assessment

For each of the aptitudes and traits, we will conduct confirmatory multi-trait multi-method analysis (MTMM) (D. T. Campbell & Fiske, 1959; Widaman, 2010) to determine convergent and discriminant validity of our measures. Traditionally, this technique as applied is most helpful to demonstrate construct validity of the measures. Recently, however, researchers have demonstrated the criticality of employing this analysis when using data with an MTMM structure for prediction with external criterion variables (Castro-Schilo et al., 2013). Convincingly, these authors demonstrated a significant bias exists if the MTMM structure is ignored when using a trait structure to predict external criterion performance. This is the situation we have when using aptitude and traits scores obtained during game play with predicting job performance criterion scores in a cyber occupation.

4.5.4. Measurement Summary

The development of sound statically valid measures of the aptitudes and traits to be employed in enhancing the selection of cyber operators is paramount. Designing a serious game that creates an entertaining experience for the player is of no benefit if it fails to deliver a veridical assessment of the aptitudes and traits under consideration. It is essential to employ an iterative approach to construct measurement whereby constructs are operationalized, piloted, statistically assessed, and revised until a solid factor structure for each construct has emerged. Both unidimensional and bi-factor models should be examined so we can be confident as to the structure of each measure. This iterative approach should be followed for each of the nine aptitudes and traits to be measured by the serious game. The constructs then can be assessed jointly via confirmatory latent variable modeling in an MTMM structure to ensure that once combined in a joint measurement paradigm their construct validity is not negatively impacted.

4.6 Usability Analysis

A major factor in the success of any serious game (or technology for that matter) is the usability of the system. As such, it is essential to perform usability analysis on the serious game. Relying on established principles from human factors, several strategies should be employed to determine its usability.

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19

Usability is the degree to which the serious game can be played effectively, efficiently, and with satisfaction on the part of the user. It is important to evaluate the serious game throughout its developmental cycle following human factors standards for usability testing and evaluation (Amstrong, Brewer, & Steinberg, 2002; Kuhnel, Seiler, Honal, & Ifenthaler, 2017). See Charlton and O'Brien (2002) and Wickens, Lee, Liu, and Gordon (2004) for overall issues and strategies in usability assessment and human factors. Based on the outcome of the evaluations, aspects of the game should be modified as necessary to ensure acceptable levels of usability are achieved.

While usability testing and evaluation is primarily employed to determine the extent to which the serious game functions within the range of human psychomotor and cognitive capabilities, special attention also needs to be paid to user satisfaction. Recently, Baek and Yoo (2018) demonstrated a holistic approach is perhaps optimal for measuring one's engagement and satisfaction. Their 13-item scale measures five factors of serious game usability: user-friendliness, personalization, speed, fun, and omnipresence. Evaluation of scores on the 13 individual items can be used to fine-tune engagement and satisfaction.

To ensure usability of the serious game, all serious game functionality and human interaction with the serious game should be identified and described following principles of job analysis from both the user-oriented and task/cognitive-oriented perspective (Coovert, 2013; Gordon, Coovert, & Elliott, 2012). The goal for usability testing is 95% of users throughout pilot testing should be able interact successfully with the serious game. Additionally, reported satisfaction with the serious game must, on average, exceed "moderately satisfied" trending toward "highly satisfied" (on a 6-point scale; 6 = highly satisfied, 5 = moderately satisfied, 4 = somewhat satisfied, 3 = somewhat dissatisfied, 2 = moderately dissatisfied, 1 = highly dissatisfied). Principles from iterative design should be followed (Gawron, Dennison, & Biferno, 2002; Wickens et al., 2004) which employs successively developing, testing, and modifying designs until usability standards are achieved.

4.7 Summary

It would be most useful to develop a serious game to further the goal of identifying and selecting individuals capable of training into one of several cyber occupations at both the enlisted and officer ranks of the USAF. We reviewed serious games; identifying common features, key elements, performance assessment, classification schemes, and engines. We argue for developing a serious game engaging to the player and one that will allow for the emergence of the aptitudes and traits that need to be measured. Limiting the number of aptitudes and traits assessed by a serious game to a manageable number, we identified nine as they are not currently assessed by other means. Depending on the scope of the game, it could cover between three and nine of those aptitudes and traits.

The serious game should employ a unifying theme and characterizing goal. Similar to the approach taken by Long and Mulch, (2017) we anticipate it could be set in a face-valid environment for cyber operations, although this is not critical. For each aptitude and trait to be assessed, developers should follow an iterative process of: developing/revising measures, pilot testing for construct dimensionality, craft/revise script and vignettes, pilot testing for usability, and revision as needed. After the sections have been completed for all measures, scripts, and

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19

vignettes they should be integrated into the full game and piloted for overall usability and validity. Revisions, as necessary, should be made. Finally, the fully-developed game should then be played by individuals in Air Force cyber AFSCs. Data can be analyzed for usability of the game and MTMM structure of the constructs. If job performance criteria are collected a concurrent validation study can be done.

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19

5.0 CONCLUSION

Examination of the literature and interviews with SMEs demonstrated there are many aptitudes and traits associated with competency in cyber occupations for both USAF enlisted and officer personnel. While many cognitive abilities and personality traits are adequately assessed with existing measures, there is a gap for some aptitudes and traits related to competency in cyber occupations. Five traits (Analytical Thinking, Adaptability, Dependability, Persistence, and Situational Awareness) and four aptitudes (Active Learning, Decision Making, Deductive and Reasoning, Systems Thinking) appear especially appropriate for further assessment. Serious games provide both a face valid and technical capability to assess these constructs. Technologically-enhanced assessment of traits and aptitudes via a serious game is an appropriate direction to take in order to enhance the selection of individuals into cyber occupations.

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19

6.0 REFERENCES

Agee, R. C., Shore, C. W., Alley, W. E., Barto, E., & Halper, M. (2009). *Air Force Officer Selection Technical Requirements Survey (AFOSTRS) Volume I; Analysis of Quantitative Results*. Randolph AFB, Texas: Air Force Personnel Center.

Air Force Personnel Center (2015). Air Force Officer Qualifying Test (AFOQT) Information Pamphlet.

Air Force Personnel Center. (2017a). *Air Force Enlisted Classification Directory (AFECD)* (p. 361).

Air Force Personnel Center. (2017b). *Air Force Officer Classification Directory (AFOCD)* (p. 262).

Air Force Personnel Center (2017c). *Evaluating Predictors for Success in Cyber (17D)*.

Air Force Personnel Center (2017d). *Evaluation of 1N4X1A/B Entry Standards*.

Alcaniz, M., Parra, E., & Giglioli, I. A. C. (2018). Virtual Reality as an Emerging Methodology for Leadership Assessment and Training. *Frontiers in Psychology*, *9*, 7. https://doi.org/10.3389/fpsyg.2018.01658

Alley, W. E., Gould, R. B., Kyllonen, P. C., Piña, J. M., Shore, C. W., Skinner, J., … Treat, B. R. (2007). Air Force Personnel Research Issues. In J. Skinner Thompson, N. (Ed.). Randolph AFB, TX: Air Force Personnel Center.

Allinson, C. W., & Hayes, J. (1996). The cognitive style index: A measure of intuition-analysis for organizational research. *Journal of Management Studies*, *33*(1), 119–135. https://doi.org/10/bn39c7

Amstrong, S. D., Brewer, W. C., & Steinberg, R. K. (2002). Usability testing. *Handbook of Human Factors Testing and Evaluation*, 403–432.

Aronson, D. (1996). Overview of systems thinking, *8*(2009), 518–560.

Baddeley. (1986). Oxford psychology series, No. 11. Working memory.

Baddeley. (2007). *Working memory, thought, and action* (Vol. 45). New York: Oxford University Press.

Baddeley, & Hitch. (1974). Working memory. *Psychology of Learning and Motivation*, *8*, 47–89. https://doi.org/10/d4vgnk

Baek, T. H., & Yoo, C. Y. (2018). Branded App Usability: Conceptualization, Measurement, and Prediction of Consumer Loyalty. *Journal of Advertising*, *47*(1), 70–82. https://doi.org/10/gftgqm

Bar-On, R. (1997). *BarOn emotional quotient inventory*. Multi-Health Systems.

Barr, A. S. (1940). Chapter II: Measurement of Teaching Ability. *Review of Educational Research*, *10*, 182–184.

Barrick, M. R., & Mount, M. K. (1991). The big five personality dimensions and job performance: a meta-analysis. *Personnel Psychology*, *44*(1), 1–26. https://doi.org/10/hjv

Barron, L. G., & Rose, M. R. (2017). Multitasking as a Predictor of Pilot Performance: Validity Beyond Serial Single-Task Assessments. *Military Psychology*, *29*(4), 316–326. https://doi.org/10/gfgzbf

Bechara, A., Damasio, H., & Damasio, A. R. (2000). Emotion, decision making and the orbitofrontal cortex. *Cerebral Cortex*, *10*(3), 295–307. https://doi.org/10/dfjr37

Bennett, R. E., & Rock, D. A. (1995). Generalizability, validity, and examinee perceptions of a computer-delivered formulating-hypotheses test. *Journal of Educational Measurement*, *32*(1), 19–36. https://doi.org/10/bhp8p8

Bhaskar, R. (2010). *Reclaiming reality: A critical introduction to contemporary philosophy*. Routledge.
Bollen, K., & Lennox, R. (1991). Conventional wisdom on measurement: A structural equation perspective. *Psychological Bulletin*, *110*(2), 305–314. https://doi.org/10/cpdtn2
Bonnechère, B., Fabris, C., Bier, J., Jan, S. V. S., Feipel, V., & Jansen, B. (2016). Evaluation of cognitive functions of aged patients using video games. In *Proceedings of the 4th Workshop on ICTs for improving Patients Rehabilitation Research Techniques* (pp. 21–24). Lisbon, Portugal: ACM.
Bottiroli, S., Tassorelli, C., Lamonica, M., Zucchella, C., Cavallini, E., Bernini, S., … Sandrini, G. (2017). Smart Aging Platform for Evaluating Cognitive Functions in Aging: A Comparison with the MoCA in a Normal Population. *Frontiers in Aging Neuroscience*, *9*. https://doi.org/10/gfkh2d
Bououd, I., & Boughzala, I. (2013). A serious game supporting collaboration. In *Proceedings of the 2013 conference on Computer supported cooperative work companion - CSCW '13* (p. 115). San Antonio, Texas: ACM Press. https://doi.org/10.1145/2441955.2441986
Brannick, M. T., Pearlman, K., & Sanchez, J. I. (2017). Work Analysis. In J. L. Farr & N. T. Tippins (Eds.), *Handbook of employee selection* (pp. 134–162). Taylor & Francis.
Bruine de Bruin, W., Parker, A. M., & Fischhoff, B. (2007). Individual differences in adult decision-making competence. *Journal of Personality and Social Psychology*, *92*(5), 938. https://doi.org/10/bgkk9z
Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, *56*(2), 81. https://doi.org/10/ds75wg
Campbell, S. G., O'Rourke, P., & Bunting, M. F. (2015). Identifying Dimensions of Cyber Aptitude: The Design of the Cyber Aptitude and Talent Assessment. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *59*(1), 721–725. https://doi.org/10/gfgzbk
Carretta, T. R., Perry, D. C., & Ree, M. J. (1996). Prediction of situational awareness in F-15 pilots. *The International Journal of Aviation Psychology*, *6*(1), 21–41. https://doi.org/10/csrd9k
Carretta, T. R., Rose, M. R., & Trent, J. D. (2016). *Air Force Officer Qualifying Test Form T: Initial Item-, Test-, Factor-, and Composite-Level Analyses*. Wright-Patterson AFB, Ohio: Air Force Research Laboratory.
Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge University Press.
Casey, E. P., Montijo, G. A., Kaiser, D., Spiker, V. A., Nullmeyer, R., Eidman, C., … Schreiber, B. T. (2008). *Air Force Research Laboratory Warfighter Readiness Research Division Participation in the 2008 IITSEC*. Mesa, AZ: Air Force Research Laboratory.
Castro-Schilo, L., Widaman, K. F., & Grimm, K. J. (2013). Neglect the structure of multitrait-multimethod data at your peril: implications for associations with external variables. *Structural Equation Modeling: A Multidisciplinary Journal*, *20*(2), 181–207. https://doi.org/10/gftckk
Chapelle, C. A. (2011). The TOEFL validity argument. In *Building a validity argument for the Test of English as a Foreign Language*^TM^ (pp. 333–366). Routledge.
Charlton, S. G., & O'Brien, T. G. (Eds.). (2002). *Handbook of human factors testing and evaluation.* Mahwah, NJ: Lawrence Erlbaum Associates.

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19

Chiaburu, D. S., Oh, I.-S., Berry, C. M., Li, N., & Gardner, R. G. (2011). The five-factor model of personality traits and organizational citizenship behaviors: a meta-analysis. *Journal of Applied Psychology*, *96*, 1140. https://doi.org/10/fnfd2q

Cobb, S. (2016). Mind this gap: criminal hacking and the global cybersecurity skills shortage, a critical analysis. Presented at the Virus Bulletin: Covering the global threat landscape, Denver, CO.

Comrey, A. L., & Lee, H. B. (2013). *A first course in factor analysis*. Psychology Press.

Coovert, M. D. (2011). Quantitative Techniques for Assessing Theory and Behavior in Continuous Learning and Collaborative Distributed Environments. In W. Bennett (Ed.), *Continuous learning in collaborative distributed environments*. Mesa, AZ: Air Force Research Laboratory.

Coovert, M. D. (2013, March 12). Petri Nets: Modeling the Complexity of Modern Jobs. https://doi.org/10.4324/9780203585146-21

Coovert, M. D., & McNelis, K. (1988). Determining the number of common factors in factor analysis: A review and program. *Educational and Psychological Measurement*, *48*(3), 687–692. https://doi.org/10/c4n67f

Coovert, M. D., Winner, J., Bennett Jr., W., & Howard, D. J. (2017). Serious Games are a Serious Tool for Team Research. *International Journal of Serious Games*, *4*(1), 41. https://doi.org/10.17083/ijsg.v4i1.141

Coovert, M. D., Winner, J., & Bennett, W. (2017). Construct Development and Validation in Game-Based Research. *Simulation & Gaming*, *48*(2), 236. https://doi.org/10.1177/1046878116682661

Credé, M., Tynan, M. C., & Harms, P. D. (2017). Much ado about grit: A meta-analytic synthesis of the grit literature. *Journal of Personality and Social Psychology*, *113*(3), 492–511. https://doi.org/10/gbtxbb

Cropley, A. (2006). In praise of convergent thinking. *Creativity Research Journal*, *18*(3), 391–404. https://doi.org/10/cn7bz9

Darr, W. (2011). Military Personality Research: A Meta-Analysis of the Self Description Inventory. *Military Psychology*, *23*, 272–296. https://doi.org/10/ch7pn6

Davis, A. C., & Stroink, M. L. (2016). The relationship between systems thinking and the new ecological paradigm. *Systems Research and Behavioral Science*, *33*(4), 575–586. https://doi.org/10/gfn5hg

Dawson, J., & Thomson, R. (2018). The Future Cybersecurity Workforce: Going Beyond Technical Skills for Successful Cyber Performance. *Frontiers in Psychology*, *9*. https://doi.org/10/gdqrm3

DiCerbo, K. E. (2014). Game-Based Assessment of Persistence. *Educational Technology & Society*, *17*, 17–28.

Dilchert, S. (2008). *Measurement and prediction of creativity at work* (Unpublished Dissertation). University of Minnesota, Minneapolis, MN.

Dörner, D., Kreuzig, H. W., Reither, F., & Stäudel, T. (1983). Lohhausen: Vom Umgang mit Unbestimmtheit und Komplexität.

Dörner, R., Göbel, S., Effelsberg, W., & Wiemeyer, J. (2016). *Serious games: foundations, concepts and practice*. Springer.

Drasgow, F., Stark, S., Chernyshenko, O. S., Nye, C. D., Hulin, C. L., & White, L. A. (2012). *Development of the tailored adaptive personality assessment system (TAPAS) to support* *army personnel selection and classification decisions*. DRASGOW CONSULTING GROUP URBANA IL.

Dreibelbis, R. C., Martin, J., Coovert, M. D., & Dorsey, D. W. (2018). The Looming Cybersecurity Crisis and What It Means for the Practice of Industrial and Organizational Psychology. *Industrial and Organizational Psychology*, *11*(02), 346–365. https://doi.org/10/gdss6p

Duckworth, A. L., Peterson, C., Matthews, M. D., & Kelly, D. R. (2007). Grit: perseverance and passion for long-term goals. *Journal of Personality and Social Psychology*, *92*(6), 1087–1101. https://doi.org/10/d2dc9t

Duin, H., & Hauge, J. B. (2008). *RefQuest: A multiplayer on-line game to support idea creation in innovation processes*.

Duin, H., Hauge, J. B., Thoben, K., & Bierwolf, R. (2009). A concept for serious gaming to support disruptive idea generation. In *2009 IEEE International Technology Management Conference (ICE)* (pp. 1–9). https://doi.org/10.1109/ITMC.2009.7461418

DuVernet, A. M., Dierdorff, E. C., & Wilson, M. A. (2015). Exploring factors that influence work analysis data: A meta-analysis of design choices, purposes, and organizational context. *Journal of Applied Psychology*, *100*(5), 1603–1633. https://doi.org/10/f7rhtk

Earnest, D. R., Allen, D. G., & Landis, R. S. (2011). Mechanisms Linking Realistic Job Previews with Turnover: A Meta-Analytic Path Analysis. *Personnel Psychology*, *64*(4), 865–897. https://doi.org/10/fbtxhg

Ehrlich, K., & Johnson-Laird, P. N. (1982). Spatial descriptions and referential continuity. *Journal of Verbal Learning and Verbal Behavior*, *21*(3), 296–306. https://doi.org/10/dfwp6j

Elsmore, T. F. (1994). SYNWORK1: A PC-based tool for assessment of performance in a simulated work environment. *Behavior Research Methods, Instruments, & Computers*, *26*(4), 421–426. https://doi.org/10/d9pkwv

Epstein, S., Pacini, R., Denes-Raj, V., & Heier, H. (1996). Individual differences in intuitive–experiential and analytical–rational thinking styles. *Journal of Personality and Social Psychology*, *71*, 390–405. https://doi.org/10/dp4gjh

Facione, P. A. (1990). Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction (The Delphi Report).

Facione, P. A., Facione, N. C., Blohm, S. W., & Giancarlo, C. A. F. (2002). *The California critical thinking skills test: CCTST*. California Academic Press San Jose, CA.

Fairbanks, B., Kucinkas, S., Nakasone, R., Trent, L., & Welsh, J. (1989). Annotated Bibliography of Armed Services Vocational Aptitude Battery (ASVAB) Validity Studies, p.1-144.

Fernandez-Aranda, F., Jimenez-Murcia, S., Santamaria, J. J., Gunnard, K., Soto, A., Kalapanidas, E., … Penelo, E. (2012). Video games as a complementary therapy tool in mental disorders: PlayMancer, a European multicentre study. *Journal of Mental Health*, *21*, 364–374. https://doi.org/10.3109/09638237.2012.664302

Fetzer, M. (2015). Serious games for talent selection and development. *TIP: The Industrial-Organizational Psychologist*, *52*, 117–125.

Fleishman, E. A., & Reilly, M. E. (1992). *Handbook of human abilities: Definitions, measurements, and job task requirements*. Consulting Psychologists Press.

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19

Foldes, H. J., Duehr, E. E., & Ones, D. S. (2008). Group differences in personality: Meta-analyses comparing five US racial groups. *Personnel Psychology*, *61*(3), 579–616. https://doi.org/10/b3q6pn

Ford, J. K., & Meyer, T. (2014). Advances in training technology: Meeting the workplace challenges of talent development, deep specialization and collaborative learning. *The Psychology of Workplace Technology*, 43–76.

Frederiksen, N., & Ward, W. C. (1978). Measures for the study of creativity in scientific problem-solving. *Applied Psychological Measurement*, *2*(1), 1–24. https://doi.org/10/bkd33f

Freese, M., & Drees, S. (2016). D-CITE - A Serious Game to Analyze Complex Decision-Making in Air Traffic Management. In A. DeGloria & R. Veltkamp (Eds.), *Games and Learning Alliance, GALA 2015, Revised Selected Papers* (Vol. 9599, pp. 23–31).

Freiknecht, J., Geiger, C., Drochtert, D., Effelsberg, W., & Dörner, R. (2016). Game Engines. In *Serious Games* (pp. 127–159). Springer.

Frensch, P. A., & Funke, J. (2014). *Complex problem solving: The European perspective*. Psychology Press.

Gael, S. (1988). *The job analysis handbook for business, industry, and government* (Vol. 1). John Wiley & Sons.

Gallagher, P., & Prestwich, S. (2012). *Supporting Cognitive Adaptability Through Game Design*.

Gawron, V. J., Dennison, T. W., & Biferno, M. A. (2002). Mock-Ups, Models, Simulations, and Embedded Testing. *Handbook of Human Factors Testing and Evaluation*, 181–223.

George, J. M., & Zhou, J. (2001). When openness to experience and conscientiousness are related to creative behavior: An interactional approach. *Journal of Applied Psychology*, *86*(3), 513–524. https://doi.org/10/bs2dtm

Gibbons, R. D., & Hedeker, D. R. (1992). Full-information item bi-factor analysis. *Psychometrika*, *57*(3), 423–436. https://doi.org/10/b3hss2

Göbel, S. (2016). Serious games application examples. In R. Dorner, S. Gobel, W. Effelsberg, & J. Wiemeyer (Eds.), *Serious Games* (pp. 319–405). Switzerland: Springer International Publishing.

Gonzalez, O., & MacKinnon, D. P. (2018). A bifactor approach to model multifaceted constructs in statistical mediation analysis. *Educational and Psychological Measurement*, *78*(1), 5–31. https://doi.org/10/gftckj

Gordon, T. G., Coovert, M. D., & Elliott, L. R. (2012). Integrating cognitive task analysis and verbal protocol analysis: A typology for describing jobs. *The Handbook of Work Analysis: Methods, Systems, Applications, and Science of Work Measurement in Organizations. New York, NY: Routledge Academic*.

Grant, D. A., & Berg, E. (1948). A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a Weigl-type card-sorting problem. *Journal of Experimental Psychology*, *38*(4), 404–411. https://doi.org/10/ft3ww5

Greiff, S., Stadler, M., Sonnleitner, P., Wolff, C., & Martin, R. (2015). Sometimes less is more: Comparing the validity of complex problem solving measures. *Intelligence*, *50*, 100–113. https://doi.org/10/f7gr9b

Hambrick, D. Z., Rench, T. A., Poposki, E. M., Darowski, E. S., Roland, D., Bearden, R. M., … Brou, R. (2011). The relationship between the ASVAB and multitasking in Navy sailors: A process-specific approach. *Military Psychology*, *23*(4), 365–380.

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19

Haywood, C., Palmer, P., & Ward, J. (1990). *Armed Services Vocational Aptitude Battery (ASVAB): The Equity of ASVAB Form 14 in the Prediction of High School Course Grades* (pp. 1–612). San Antonio, TX: Operational Technologies Corporation.

Hocevar, D. (1981). Measurement of creativity: Review and critique. *Journal of Personality Assessment*, *45*, 450–464. https://doi.org/10/cqf77j

Holzinger, K. J., & Swineford, F. (1939). A study in factor analysis: The stability of a bi-factor solution. *Supplementary Educational Monographs*.

Hunter, J. E., & Hunter, R. F. (1984). Validity and utility of alternative predictors of job performance. *Psychological Bulletin*, *96*, 72. https://doi.org/10/d2pj27

Hurtt, R. K. (2010). Development of a scale to measure professional skepticism. *Auditing: A Journal of Practice & Theory*, *29*, 149–171. https://doi.org/10/cpvvdc

Hurtt, R. K., Brown-Liburd, H., Earley, C. E., & Krishnamoorthy, G. (2013). Research on auditor professional skepticism: Literature synthesis and opportunities for future research. *Auditing: A Journal of Practice & Theory*, *32*(sp1), 45–97. https://doi.org/10/gfn5hh

Jalali, M. S., Siegel, M., & Madnick, S. (2017). Decision Making and Biases in Cybersecurity Capability Development: Evidence from a Simulation Game Experiment. *SSRN Electronic Journal*. https://doi.org/10/gfbzrf

James, L. R., Demaree, R. G., & Wolf, G. (1984). Estimating within-group interrater reliability with and without response bias. *Journal of Applied Psychology*, *69*(1), 85. https://doi.org/10/g5c

Jennrich, R. I., & Bentler, P. M. (2011). Exploratory bi-factor analysis. *Psychometrika*, *76*(4), 537–549. https://doi.org/10/fmhqs5

Jensen, R., Presnell, B., Lunsford, J., & Cobb, M. G. (2014). *Automated intelligent training with a tactical decision making serious game*. San Mateo, CA: Stottler Henke Associates Inc.

Johnston, M., Carpenter, A., & Hale, K. (2011). Test-Retest Reliability of CogGauge: A Cognitive Assessment Tool for SpaceFlight. In D. Harris (Ed.), *Engineering Psychology and Cognitive Ergonomics* (pp. 565–571). Springer Berlin Heidelberg.

Jose, I., LaPort, K., & Trippe, D. M. (2016). Requisite attributes for cyber security personnel and teams: Cyber risk mitigation through talent management. In *Psychosocial dynamics of cyber security* (pp. 167–198). Routledge.

Kaufman, A. (1993). Manual for the Kaufman Adolescent and Adult Intelligence Test (KAIT). *Circle Pines, MN: American Guidance Service*.

Kozlowski, S. W., & DeShon, R. P. (2004). A psychological fidelity approach to simulation-based training: Theory, research and principles. *Scaled Worlds: Development, Validation, and Applications*, 75–99.

Kuhnel, M., Seiler, L., Honal, A., & Ifenthaler, D. (2017). Mobile Learning Analytics in Higher Education: Usability Testing and Evaluation of an APP Prototype. *International Association for Development of the Information Society*.

Lambert, B. (2014a). Occupational analysis report AFSC 3D0X2 cyber systems operations OSSN 2977.

Lambert, B. (2014b). Occupational analysis report AFSC 3D0X3 cyber surety OSSN 2978.

Landers, R. N., Auer, E. M., Collmus, A. B., & Armstrong, M. B. (2018). Gamification Science, Its History and Future: Definitions and a Research Agenda. *Simulation & Gaming*, *49*(3), 315–338. https://doi.org/10.1177/1046878118774385

LeBreton, J. M., & Senter, J. L. (2008). Answers to 20 Questions About Interrater Reliability and Interrater Agreement. *Organizational Research Methods*, *11*(4), 815–852. https://doi.org/10/frx8kx

Lecznar, W. B., & Klesch, J. K. (1965). *Development and preliminary validation of the electronic data processing test-63* (No. PRL-TR-65-12) (pp. 1–15). Lackland Air Force Base, Texas: Personnel Research Laboratory.

Lewis, S., Ellis, J. B., & Kellogg, W. A. (2010). Using virtual interactions to explore leadership and collaboration in globally distributed teams. In *Proceedings of the 3rd international conference on Intercultural collaboration - ICIC '10* (p. 9). Copenhagen, Denmark: ACM Press. https://doi.org/10.1145/1841853.1841856

Lif, P., Granåsen, M., & Sommestad, T. (2017). Development and validation of technique to measure cyber situation awareness (pp. 1–8). Presented at the Cyber Situational Awareness, Data Analytics And Assessment (Cyber SA), 2017 International Conference On, IEEE.

Linehan, C., Lawson, S., Doughty, M., & Kirman, B. (2009). Developing a serious game to evaluate and train group decision making skills. In *Proceedings of the 13th International MindTrek Conference: Everyday Life in the Ubiquitous Era on - MindTrek '09* (p. 106). Tampere, Finland: ACM Press. https://doi.org/10.1145/1621841.1621861

Linkov, I., Fenton, G., Satterstrom, F. K., Gaskins, R., & Lewis, B. (2007). *Computer-Mediated Training Tools to Enhance Joint Task Force Cognitive Leadership Skills*. Alexandria, VA: Army Research Inst For the Behavioral and Social Sciences.

Long, D. T., & Mulch, C. M. (2017). *Interactive Wargaming Cyberwar: 2025*. Naval Postgraduate School Monterey United States. Retrieved from https://apps.dtic.mil/docs/citations/AD1053350

Looper, T. (1997). Changing Technology and Military Missions: Impacts on US Military Personnel Systems, p.1-6.

Ludoscience. (2018). Retrieved from http://serious.gameclassification.com/

Luthans, F. (2002). The need for and meaning of positive organizational behavior. *Journal of Organizational Behavior*, *23*, 695–706. https://doi.org/10/b28d7w

Malviya, A., Fink, G. A., Sego, L., & Endicott-Popovsky, B. (2011). Situational awareness as a measure of performance in cyber security collaborative work (pp. 937–942). Presented at the Information Technology: New Generations (ITNG), 2011 Eighth International Conference on, IEEE.

Manley, G. G. (2016). *Technical Reports on Cross-Validation of AFOQT Form S for Cyberspace Operations - Cyberspace Control*. Randolph AFB, TX: Air Force Personnel Center.

Manley, G. G., & Weissmuller, J. (2017). Development of the United States Air Force's Self-Description Inventory. Presented at the 32rd Annual Meetings of the Society for Industrial and Organizational Psychology.

Martin, M. M., & Rubin, R. B. (1995). A new measure of cognitive flexibility. *Psychological Reports*, *76*(2), 623–626. https://doi.org/10/b6chh8

Mayer, J. D., Salovey, P., Caruso, D. R., & Sitarenios, G. (2003). Measuring emotional intelligence with the MSCEIT V2. 0. *Emotion*, *3*(1), 97–105. https://doi.org/10/fpb3gp

McAllister, D. J. (1995). Affect- and Cognition-Based Trust as Foundations for Interpersonal Cooperation in Organizations. *Academy of Management Journal*, *38*(1), 24–59. https://doi.org/10/bsdqqf

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19

Mohan, D., Angus, D. C., Ricketts, D., Farris, C., Fischhoff, B., Rosengart, M. R., … Barnato, A. E. (2014, August). Assessing the Validity of Using Serious Game Technology to Analyze Physician Decision Making. *Plos One*, *9*.

Monsell, S. (1978). Recency, immediate recognition memory, and reaction time. *Cognitive Psychology*, *10*, 465–501. https://doi.org/10/cd9sk9

Moore, S., Dolansky, M., Singh, M., Palmieri, P., & Alemi, F. (2010). The systems thinking scale. *Cleveland: Case Western Reserve University*.

Morgeson, F. P., & Campion, M. A. (1997). Social and Cognitive Sources of Potential Inaccuracy in Job Analysis. *Journal of Applied Psychology*, *82*(5), 627–655. https://doi.org/10/dfw39r

Mumford, M. D., Marks, M. A., Connelly, M. S., Zaccaro, S. J., & Johnson, J. F. (1998). Domain-based scoring in divergent-thinking tests: Validation evidence in an occupational sample. *Creativity Research Journal*, *11*, 151–163. https://doi.org/10/bzgnc2

Muthén, L. K., & Muthén, B. (2015). Mplus. *The Comprehensive Modelling Program for Applied Researchers: User's Guide*, *5*.

Nasreddine, Z. S., Phillips, N. A., Bédirian, V., Charbonneau, S., Whitehead, V., Collin, I., … Chertkow, H. (2005). The Montreal Cognitive Assessment, MoCA: a brief screening tool for mild cognitive impairment. *Journal of the American Geriatrics Society*, *53*(4), 695–699. https://doi.org/10/dmt678

National Research Council. (2015). *Measuring human capabilities: An agenda for basic research on the assessment of individual and group performance potential for military accession*. National Academies Press.

Navarro, A., Pradilla, J. V., & Madriñan, P. (2010). Work in progress — Serious 3D game for mobile networks planning. In *2010 IEEE Frontiers in Education Conference (FIE)* (pp. T1F-1). https://doi.org/10.1109/FIE.2010.5673539

NeSmith, B. (2018). The Cybersecurity Talent Gap Is an Industry Crisis. Retrieved January 16, 2019, from https://www.forbes.com/sites/forbestechcouncil/2018/08/09/the-cybersecurity-talent-gap-is-an-industry-crisis/#35f6780a6b36

Nonaka, T., Miki, K., Odajima, R., & Mizuyama, H. (2016). Analysis of dynamic decision making underpinning supply chain resilience: A serious game approach. *Ifac Papersonline*, *49*, 474–479. https://doi.org/10.1016/j.ifacol.2016.10.632

Nye, C. D., Beal, S. A., Drasgow, F., Dressel, J. D., White, L. A., & Stark, S. (2014). *Assessing the tailored adaptive personality assessment system*. Urbana, IL: Drasgow Consulting Group.

O*NET Abilities. (n.d.). Retrieved August 2, 2018, from https://www.onetonline.org/find/descriptor/browse/Abilities/

O*NET Skills. (n.d.). Retrieved August 2, 2018, from https://www.onetonline.org/find/descriptor/browse/Skills/

O'Rourke, P., Karuzis, V. P., Kim, S., Tseng, A., Pandža, N. B., Young, J. M., … Campbell, S. G. (2017). Assessing aptitude for USAF cyber warfare operations training, 143.

O'Rourke, P., Karuzis, V. P., S. Kim, A. Tseng, Pandža, N. B., Young, J. M., … Campbell, S. G. (2017). *Assessing aptitude for USAF cyber warfare operations training* (pp. 1–143). College Park, MD: University of Maryland Center for Advanced Study of Language.

Paullin, C., Ingerick, M., Trippe, D. M., & Wasko, L. (2011). *Identifying Best Bet Entry-Level Selection Measures for US Air Force Remotely Piloted Aircraft (RPA) Pilot and Sensor Operator (SO) Occupations*. Randolph AFB, Texas: Air Force Personnel Center.

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19

Peterson, J. B. (2002). *Maps of meaning: The architecture of belief*. Routledge.
Peterson, N. G., Mumford, M. D., Borman, W. C., Jeanneret, P. R., Fleishman, E. A., Levin, K. Y., … Pearlman, K. (2001). Understanding work using the Occupational Information Network (O* NET): Implications for practice and research. *Personnel Psychology*, *54*(2), 451–492. https://doi.org/10/ds3d77
Pirolli, P., & Card, S. (1999). Information foraging. *Psychological Review*, *106*(4), 643. https://doi.org/10/fdtn32
Poposki, E. M., Oswald, F. L., & Brou, R. J. (2009). *Development of a new measure of polychronicity*. Michigan State Univ East Lansing.
Powers, R. B., & Kirkpatrick, K. (2013). Playing With Conflict: Teaching Conflict Resolution Through Simulations and Games. *Simulation & Gaming*, *44*(1), 51–72. https://doi.org/10/gfbzsm
Raybourn, E. M., Deagle, M. E., Mendini, K., & Heneghan, J. (2005). Adaptive Thinking & Leadership Simulation Game Training for Special Forces Officers. In *2005 Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC)* (pp. 1–9). Orlando, FL.
Raybourn, E. M., Roberts, B., Diller, D., & Dubow, L. (2008). *Honing intercultural engagement skills for stability operations with darwars ambush! nk game-based training*. Orlando, FL: Army Simulation Training and instrumentation Command.
Ree, M., Valentine, L., & Wilbourn, J. (1984). Relationships of the Armed Services Vocational Aptitude Battery (ASVAB) Forms 8, 9, and 10 to Air Force Technical School Final Grades, p. 1-175.
Reise, S. P., Waller, N. G., & Comrey, A. L. (2000). Factor analysis and scale revision. *Psychological Assessment*, *12*, 287. https://doi.org/10/bdgsn2
Resnikoff, H. L. (2012). *The illusion of reality*. Springer Science & Business Media.
Roth, P. L., Bevier, C. A., Bobko, P., Switzer, F. S., & Tyler, P. (2001). Ethnic group differences in cognitive ability in employment and educational settings: A meta-analysis. *Personnel Psychology*, *54*, 297–330. https://doi.org/10/dpds25
Russell, T. L., & Sellman, W. . (2009). *Information and Communication Technology Literacy Test Training School Validation: Phase II Final Report*. Randolph AFB, Texas: Air Force Personnel Center.
Schell, J. (2008). *The art of game design: A deck of lenses*. AK Peters/CRC Press.
Schier, S., Freese, M., & Muhlhausen, T. (2016). Serious Gaming in Airport Management: Transformation from a Validation Tool to a Learning Environment. In R. Bottino, J. Jeuring, & R. C. Veltkamp (Eds.), *Games and Learning Alliance, Gala 2016* (Vol. 10056, pp. 187–196).
Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, *124*, 262. https://doi.org/10/fd8vxp
Schweizer, F., Wüstenberg, S., & Greiff, S. (2013). Validity of the MicroDYN approach: Complex problem solving predicts school grades beyond working memory capacity. *Learning and Individual Differences*, *24*, 42–52. https://doi.org/10/f4vk74
Shute, V. J. (2011). Stealth assessment in computer-based games to support learning. *Computer Games and Instruction*, *55*, 503–524.

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19

Simpson, R. W., & Nester, M. A. (2007). Taxonomy for reasoning questions using logic based measurement. Presented at the International Public Management Association for Human Resource Conference.

Sinkula, J. M., Baker, W. E., & Noordewier, T. (1997). A framework for market-based organizational learning: Linking values, knowledge, and behavior. *Journal of the Academy of Marketing Science*, *25*, 305. https://doi.org/10/d44t7c

Skinner, J. (2006). *Criterion-related validity of the operational AFOQT composites.* San Antonio, TX: Operational Technologies Corporation.

Spiro, R. J., & Jehng, J. (1990). Cognitive Flexibility, random access instruction and hypertext: Theory and technology for the nonlinear and multi-dimensional traversal of complex subject matter. *Cognition, Education, and Multimedia: Exploring Ideas in High Technology*, 163–205.

Stark, S., Chernyshenko, O., & Drasgow, F. (2012). Constructing fake-resistant personality tests using item response theory: High stakes personality testing with multidimensional pairwise preferences. *New Perspectives on Faking in Personality Assessments; NY: Oxford University Press*, 215.

Stark, S., Chernyshenko, O. S., Drasgow, F., Nye, C. D., White, L. A., Heffner, T., & Farmer, W. L. (2014). From ABLE to TAPAS: A new generation of personality tests to support military selection and classification decisions. *Military Psychology*, *26*(3), 153–164. https://doi.org/10/gckf9v

Stark, S., Chernyshenko, O. S., Nye, C. D., Drasgow, F., & White, L. A. (2017). *Moderators of the Tailored Adaptive Personality Assessment System Validity*. Urbana, IL: Drasgow Consulting Group.

Sternberg, S. (1966). High-speed scanning in human memory. *Science*, *153*(3736), 652–654. https://doi.org/10/b4m7mh

Swezey, R. W., Streufert, S., Unger, K. W., & Van Rijn, P. (1985). *Development of a Computer Simulation for Assessing Decision-Making Style Using Cognitive Complexity Theory* (p. 86).

Teo, L., John, B. E., & Pirolli, P. (2007). Towards a tool for predicting user exploration (pp. 2687–2692). Presented at the CHI'07 Extended Abstracts on Human Factors in Computing Systems, ACM.

Thompson, N. (2007). Enlisted Selection and Classification Tests: Precursors of the ASVAB, p.1-27.

Tippins, N. T. (2015). Technology and Assessment in Selection. *Annual Review of Organizational Psychology and Organizational Behavior*, *2*(1), 551–582. https://doi.org/10/gckgb2

Toivola, J. (2018). A global market survey: 10 companies offering game-based assessment. Retrieved from http://psykometriikka.fi/kirjoitukset/a-global-market-survey-10-companies-offering-game-based-assessment

Torgerson, T. (1934). The measurement and prediction of teaching ability. *Review of Educational Research*, *4*, 261–266.

Tost, D., Pazzi, S., Barnekow, A. von, Felix, E., Puricelli, S., & Bottiroli, S. (2014). SmartAgeing: a 3D serious game for early detection of mild cognitive impairments. In *Proceedings of the 8th International Conference on Pervasive Computing Technologies for Healthcare* (pp. 294–297). Oldenburg, Germany: ICST (Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering).

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19

Trippe, D. M., Moriarty, K. O., Beatty, A. S., & Diaz, T. E. (2014). *Cyber Test Form Development and Follow-on Cyber Applications*. Randolph AFB, Texas: Air Force Personnel Center.

Trippe, D. M., & Russell, T. L. (2011). *Information and Communication Technology Literacy Test Norming Study: Phase III Final Report*. Randolph AFB, Texas: Air Force Personnel Center.

Trippe, D. M., Russell, T. L., Schwartz, K. L., & Weissmuller, J. J. (2008). *Issues in Information and Communication Technology Test Development: A literature Review and Summary of Best Practice*. Randolph AFB, Texas: Air Force Personnel Center.

Tsarava, K., Moeller, K., & Ninaus, M. (2018). Training Computational Thinking through board games: The case of Crabs & Turtles. *International Journal of Serious Games*, *5*(2), 25. https://doi.org/10.17083/ijsg.v5i2.248

Turner, J. H., & Valentine, S. R. (2001). Cynicism as a fundamental dimension of moral decision-making: A scale development. *Journal of Business Ethics*, *34*(2), 123–136. https://doi.org/10/fqwgg2

Vegt, N., Visch, V., Vermeeren, A., & de Ridder, H. (2016). Player Experiences and Behaviors in a Multiplayer Game: designing game rules to change interdependent behavior. *International Journal of Serious Games*, *3*(4), 69. https://doi.org/10.17083/ijsg.v3i4.150

Ventura, M., & Shute, V. (2013). The validity of a game-based assessment of persistence. *Computers in Human Behavior*, *29*, 2568–2572. https://doi.org/10.1016/j.chb.2013.06.033

Wear, H. E. (2018). *A Validation of a Game-Based Assessment for The Measurement of Vocational Interest*. San Francisco State University.

Welsh, J. (1997). Alternate Specifications for the Armed Services Vocational Aptitude Battery (ASVAB): Subgroup Differences Study based on the Exposure to Content Approach, p/1-5.

Whalen, K. A., Berlin, C., Ekberg, J., Barletta, I., & Hammersberg, P. (2018). “All they do is win”: Lessons learned from use of a serious game for Circular Economy education. *Resources Conservation and Recycling*, *135*, 335–345. https://doi.org/10.1016/j.resconrec.2017.06.021

Whalen, K., & Peck, D. (2014). In the Loop–Sustainable, Circular Product Design and Critical Materials. *International Journal of Automation Technology*, *8*(5), 664–676. https://doi.org/10/gfs7tv

Wickens, C. D., Lee, J. D., Liu, Y., & Gordon-Becker, S. (2004). Cognition. *An Introduction to Human Factors Engineering*, *2*, 120–155.

Widaman, K. F. (2010). Multitrait-multimethod analysis. *The Reviewer’s Guide to Quantitative Methods in the Social Sciences*, 299–314.

Wiemeyer, J., & Hardy, S. (2013). Serious games and motor learning: concepts, evidence, technology. In *Serious Games and Virtual Worlds in Education, Professional Development, and Healthcare* (pp. 197–220). IGI Global.

Wiemeyer, J., & Kliem, A. (2012). Serious games in prevention and rehabilitation—a new panacea for elderly people? *European Review of Aging and Physical Activity*, *9*(1), 41. https://doi.org/10/bfxfmn

Williams, B., Bates, B., McGovern, T., Heuring, T., & Walton, G. (2006). *MS&G, When Worlds Collide: A Primer for Potential*. Alexandria, VA: Institute for Defense Analyses.

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19

Zwaan, R. A., & Radvansky, G. A. (1998). Situation models in language comprehension and memory. *Psychological Bulletin*, *123*, 162–185. https://doi.org/10/cntx9t

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19

APPENDIX A

Complete List of Cyber Aptitudes and Traits with Definitions

Aptitude	Definition
Active Learning	The ability to understand the implications of new information for both current and future problem-solving and decision-making.
Anomaly detection	The ability to detect information that is anomalous in a larger context, such that it does not conform to the expected pattern.
Category Flexibility	The ability to generate or use different sets of rules for combining or grouping things in different ways.
Complex problem-solving	The ability to learn and effectively manipulate systems, which are complex, opaque, and dynamic
Convergent Creative Thinking	The ability to explore a variety of solutions by forming connections between concepts that are typically weakly related or unrelated, and ultimately to hone in on one single, correct solution to a problem.
Decision Making	The ability to consider the relative costs and benefits of potential actions to choose the most appropriate one.
Deductive Reasoning	The ability to apply general rules to specific problems to produce answers that make sense.
Emotional Intelligence	The ability to be aware of others' reactions and understand why they react as they do
Flexibility of closure	The ability to identify or detect a known pattern (a figure, object, word, or sound) that is hidden in other distracting material.
Fluency of Ideas	The ability to come up with a number of ideas about a topic (the number of ideas is important, not their quality, correctness, or creativity).
Inductive Reasoning	The ability to combine pieces of information to form general rules or conclusions (includes finding a relationship among seemingly unrelated events).
Information and Technology Aptitude	The ability to use a computer, communication devices, and related applications to input, retrieve, and communicate information.
Information Ordering	The ability to arrange things or actions in a certain order or pattern according to a specific rule or set of rules (e.g., patterns of numbers, letters, words, pictures, mathematical operations).
Instructing	The ability to teach others how to do something.
Mathematical Reasoning	The ability to choose the right mathematical methods or formulas to solve a problem.
Memorization	The ability to remember information such as words, numbers, pictures, and procedures.
Mental Agility	The ability to restructure one's knowledge as an adaptive response to changing situational demands
Mental Model Ability	The ability to construct abstract, internal representations of a situation, real or imagined, derived from a narrative or other form of input and

	provide a basis for inference making and successful recall of information.
Modeling program execution	The ability to scan incoming information, detect patterns and react quickly.
Number Facility	The ability to add, subtract, multiply, or divide quickly and correctly.
Oral Comprehension	The ability to listen to and understand information and ideas presented through spoken words and sentences.
Oral Expression	The ability to communicate information and ideas in speaking so others will understand.
Originality	The ability to come up with unusual or clever ideas about a given topic or situation, or to develop creative ways to solve a problem.
Pattern recognition and scanning	The ability to determine the rules that govern a pattern.
Perceptual Speed	The ability to quickly and accurately compare similarities and differences among sets of letters, numbers, objects, pictures, or patterns. The things to be compared may be presented at the same time or one after the other. This ability also includes comparing a presented object with a remembered object.
Problem Sensitivity	The ability to tell when something is wrong or is likely to go wrong. It does not involve solving the problem, only recognizing there is a problem.
Psychomotor Speed	The ability to respond quickly and accurately in the face of proactive interference.
Resistance to Interference	The ability to respond quickly and to control the speeded motor response in the face of interference.
Selective Attention	The ability to concentrate on a task over a period of time without being distracted.
Spatial Visualization	The ability to form and manipulate visuospatial representations.
Speed of Closure	The ability to quickly make sense of, combine, and organize information into meaningful patterns.
Systems Thinking	The ability to understand of how multiple parts of a system interact and influence each other.
Time Sharing	The ability to shift back and forth between two or more activities or sources of information (such as speech, sounds, touch, or other sources).
Time management	The ability to manage one's own time and the time of others.
Vigilance	The ability to remain vigilant or sustain attention during a task that occurs over a prolonged period of time.
Visuospatial working memory	The ability to briefly hold and manipulate information from the spatial domain.
Written Comprehension	The ability to read and understand information and ideas presented in writing.
Written Expression	The ability to communicate information and ideas in writing so others will understand.

Traits	Definition
Achievement	The degree to which individuals establish and maintain personally challenging achievement goals and exerting effort toward mastering tasks.
Adaptability	The degree to which individuals are open to change (positive or negative) and to considerable variety in the workplace.
Analytical Thinking	The degree to which individuals analyze information and use logic to address work-related issues and problems.
Attention to Detail	The degree to which individuals are careful about detail and thorough in completing work tasks.
Concern for Others	The degree to which individuals are sensitive to others' needs and feelings and being understanding and helpful on the job.
Cooperation	The degree to which individuals are pleasant with others on the job and display a good-natured, cooperative attitude.
Dependability	The degree to which individuals are reliable, responsible, and dependable, and fulfill obligations.
Independence	The degree to which individuals develop their own ways of doing things, guide themselves with little or no supervision, and depend on themselves to get things done.
Initiative	The degree to which individuals have willingness to take on responsibilities and challenges.
Innovation	The degree to which individuals use creativity and alternative thinking to develop new ideas for and answers to work-related problems.
Integrity	The degree to which individuals are honest and ethical.
Leadership	The degree to which individuals have a willingness to lead, take charge, and offer opinions and direction.
Need for Cognition	The degree to which individuals enjoy participating in mentally demanding tasks.
Need for Cognitive Closure	The degree to which individuals have the need to arrive at a solution during problem solving
Persistence	The degree to which individuals have persistence in the face of obstacles.
Resilience	The degree to which individuals have the capacity to rebound or bounce back from adversity, conflict, and failure or even positive events, progress, and increased responsibility.
Self-Control	The degree to which individuals maintain composure, keep emotions in check, control anger, and avoid aggressive behavior, even in very difficult situations.
Self-Discipline	The degree to which individuals tend be focused and dedicated to working hard and completing tasks in a timely manner.
Situational Awareness	The degree to which individuals pay attention to their surroundings and rarely get lost or surprised.
Skepticism	The degree to which individuals distrust others.
Social Orientation	The degree to which individuals prefer to work with others rather than alone, and are personally connected with others on the job.

Distribution A: Approved for public release. 88 ABW-2019-xxxx, cleared ??

Stress Tolerance	The degree to which individuals accept criticism and deal calmly and effectively with high stress situations.
Team Player	The degree to which individuals prefer working in teams and help people work together better.
Tolerance for Risk	The degree to which individuals are risk taking.

Distribution A: Approved for public release. 88 ABW-2019-xxxx, cleared ??

APPENDIX B

SME Recruitment Email for Job Analysis Survey

Hello [NAME OF RECIPIENT],

My name is Jackie Martin and I am one of the graduate students working under Mike Coovert at USF on the Cyber Assessment Project with the Air Force. Thank you for the great feedback on the project that you provided during our phone call in April!

You indicated at the end of the call that you were open to further supporting the project, so we would love for you to provide further SME input in a brief (15-30 minutes) survey designed to further refine the important aptitude and traits for cyber positions in the Air Force.

If you are able to assist with this, please find the link to the online version of the survey below. If you'd prefer to print the survey to complete and email back, a PDF of the survey is also attached. Additionally, if you know of anyone else who would be able to provide SME input for cyber positions in the Air Force, we would much appreciate you forwarding the survey to them or referring us to them to contact directly.

https://usf.az1.qualtrics.com/jfe/form/SV_23MwcrBVzetfQRT

Thank you for your time and please let me know if you have any questions.

Thank you,
Jackie

APPENDIX C

SME Survey

SECTION 1: BACKGROUND INFORMATION

1. **What is your AFSC** *(e.g., 3D013)*

2. **What is your current position?** *(e.g., Cyber Surety Specialist)*

Specify: ____**--FOR THE REMAINDER OF THE SURVEY, WHEN WE ASK YOU ABOUT "YOUR AFS" THIS IS WHAT WE ARE REFERRING TO**

3. **How long have you been in your current position?**

4. **How typical is your current position compared to what most people in your AFSC generally do?**
   - ☐ **Very typical** (most jobs are like mine)
   - ☐ **Somewhat typical** (about half of the jobs are like mine)
   - ☐ **Not that typical** (fewer than a quarter of the jobs are like mine)
   - ☐ **Completely different** (nearly all the jobs are different than mine)

5. **What year were you born?** Year: ___________

6. **Are you male or female?**
   - ☐ Male
   - ☐ Female
   - ☐ Other *(please specify)* ___________
   - ☐ Prefer not to answer

7. **Indicate the highest level of education that you have completed.** *(Mark one)*
   - ☐ **Less than a High School Diploma**
   - ☐ **High School Diploma** (or GED)
   - ☐ **Some College Courses**
   - ☐ **Associate's Degree** (or other 2-year degree)
   - ☐ **Bachelor's Degree**
   - ☐ **Post-Graduate Degree** (Master's, MBA, law, MD, PhD, etc.)

8. **Do you have any cyber related certifications?**
   - ☐ No
   - ☐ Yes *(please list relevant certifications)*

____

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19

SECTION 2: APTITUDES

Please review the list of aptitudes (enduring attributes that influence performance) provided below:

Aptitude	Definition
Active learning	The ability to understand the implications of new information for both current and future problem solving and decision-making.
Anomaly detection	The ability to detect information that is anomalous in a larger context, such that it does not conform to the expected pattern.
Category flexibility	The ability to generate or use different sets of rules for combining or grouping things in different ways.
Complex problem-solving	The ability to learn and effectively manipulate systems, which are complex, opaque, and dynamic
Convergent creative thinking	The ability to explore a variety of solutions by forming connections between concepts that are typically weakly related or unrelated, and ultimately to hone in on one single, correct solution to a problem.
Decision making	The ability to consider the relative costs and benefits of potential actions to choose the most appropriate one.
Deductive reasoning	The ability to apply general rules to specific problems to produce answers that make sense.
Emotional intelligence	The ability to be aware of others' reactions and understand why they react as they do
Flexibility of closure	The ability to identify or detect a known pattern (a figure, object, word, or sound) that is hidden in other distracting material.
Fluency of ideas	The ability to come up with a number of ideas about a topic (the number of ideas is important, not their quality, correctness, or creativity).
Inductive reasoning	The ability to combine pieces of information to form general rules or conclusions (includes finding a relationship among seemingly unrelated events).
Information and technology aptitude	The ability to use a computer, communication devices, and related applications to input, retrieve, and communicate information.
Information ordering	The ability to arrange things or actions in a certain order or pattern according to a specific rule or set of rules (e.g., patterns of numbers, letters, words, pictures, mathematical operations).
Instructing	The ability to teach others how to do something.
Mathematical reasoning	The ability to choose the right mathematical methods or formulas to solve a problem.
Memorization	The ability to remember information such as words, numbers, pictures, and procedures.

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19

Aptitude	Definition
Mental agility	The ability to restructure one's knowledge as an adaptive response to changing situational demands
Mental model ability	The ability to construct abstract, internal representations of a situation, real or imagined, derived from a narrative or other form of input and provide a basis for inference making and successful recall of information.
Modeling program execution	The ability to scan incoming information, detect patterns and react quickly.
Number facility	The ability to add, subtract, multiply, or divide quickly and correctly.
Oral comprehension	The ability to listen to and understand information and ideas presented through spoken words and sentences.
Oral expression	The ability to communicate information and ideas in speaking so others will understand.
Originality	The ability to come up with unusual or clever ideas about a given topic or situation, or to develop creative ways to solve a problem.
Pattern recognition and scanning	The ability to determine the rules that govern a pattern.
Perceptual speed	The ability to quickly and accurately compare similarities and differences among sets of letters, numbers, objects, pictures, or patterns. The things to be compared may be presented at the same time or one after the other. This ability also includes comparing a presented object with a remembered object.
Problem sensitivity	The ability to tell when something is wrong or is likely to go wrong. It does not involve solving the problem, only recognizing there is a problem.
Psychomotor speed	The ability to respond quickly and accurately in the face of proactive interference.
Resistance to interference	The ability to respond quickly and to control the speeded motor response in the face of interference.
Selective attention	The ability to concentrate on a task over a period of time without being distracted.
Spatial visualization	The ability to form and manipulate visuospatial representations.
Speed of closure	The ability to quickly make sense of, combine, and organize information into meaningful patterns.
Systems thinking	The ability to understand of how multiple parts of a system interact and influence each other.
Time sharing	The ability to shift back and forth between two or more activities or sources of information (such as speech, sounds, touch, or other sources).
Time management	The ability to manage one's time and the time of others.
Vigilance	The ability to remain vigilant or sustain attention during a task that occurs over a prolonged period.

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19

Aptitude	Definition
Visuospatial working memory	The ability to briefly hold and manipulate information from the spatial domain.
Written comprehension	The ability to read and understand information and ideas presented in writing.
Written expression	The ability to communicate information and ideas in writing so others will understand.

9. **Can you think of any aptitudes that are important for your position that are missing from this list?**
   - ☐ **No**
   - ☐ **Yes**

If you selected yes, please name the aptitude and briefly describe it.

Aptitude Name: ____________________________________

Description:

____

____

____

____

____

____

____

____

____

____

____

____

____

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19

PART 1 - Importance:

10. Using a scale from 1 to 5, how important is each aptitude to the performance in the position you indicated at the beginning of this survey. Importance refers to the significance of an ability for accomplishing one's job responsibilities.

PART 2 - Needed at entry:

11. Indicate whether is each aptitude is needed at entry (upon starting the position) or not needed at entry (can be learned or trained while in the position).

	PART 1: IMPORTANCE					PART 2: NEEDED AT ENTRY	
APTITUDE:	**Not at all important**	**Slightly Important**	**Moderately Important**	**Very Important**	**Extremely Important**	**Not required at entry**	**Required at entry**
a. Active learning	1	2	3	4	5	No	Yes
b. Anomaly detection	1	2	3	4	5	No	Yes
c. Category flexibility	1	2	3	4	5	No	Yes
d. Complex problem-solving	1	2	3	4	5	No	Yes
e. Convergent creative thinking	1	2	3	4	5	No	Yes
f. Decision making	1	2	3	4	5	No	Yes
g. Deductive reasoning	1	2	3	4	5	No	Yes
h. Emotional intelligence	1	2	3	4	5	No	Yes
i. Flexibility of closure	1	2	3	4	5	No	Yes
j. Fluency of ideas	1	2	3	4	5	No	Yes
k. Inductive reasoning	1	2	3	4	5	No	Yes
l. Information and technology aptitude	1	2	3	4	5	No	Yes
m. Information ordering	1	2	3	4	5	No	Yes
n. Instructing	1	2	3	4	5	No	Yes
o. Mathematical reasoning	1	2	3	4	5	No	Yes
p. Memorization	1	2	3	4	5	No	Yes
q. Mental agility	1	2	3	4	5	No	Yes
r. Mental model ability	1	2	3	4	5	No	Yes
s. Modeling program execution	1	2	3	4	5	No	Yes
t. Number facility	1	2	3	4	5	No	Yes
u. Oral comprehension	1	2	3	4	5	No	Yes

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19

	PART 1: IMPORTANCE					PART 2: NEEDED AT ENTRY	
APTITUDE:	**Not at all important**	**Slightly Important**	**Moderately Important**	**Very Important**	**Extremely Important**	**Not required at entry**	**Required at entry**
v. Oral expression	1	2	3	4	5	No	Yes
w. Originality	1	2	3	4	5	No	Yes
x. Pattern recognition and scanning	1	2	3	4	5	No	Yes
y. Perceptual speed	1	2	3	4	5	No	Yes
z. Problem sensitivity	1	2	3	4	5	No	Yes
aa. Psychomotor speed	1	2	3	4	5	No	Yes
bb. Resistance to interference	1	2	3	4	5	No	Yes
cc. Selective attention	1	2	3	4	5	No	Yes
dd. Spatial visualization	1	2	3	4	5	No	Yes
ee. Speed of closure	1	2	3	4	5	No	Yes
ff. Systems thinking	1	2	3	4	5	No	Yes
gg. Time sharing	1	2	3	4	5	No	Yes
hh. Time management	1	2	3	4	5	No	Yes
ii. Vigilance	1	2	3	4	5	No	Yes
jj. Visuospatial working memory	1	2	3	4	5	No	Yes
kk. Written comprehension	1	2	3	4	5	No	Yes
ll. Written expression	1	2	3	4	5	No	Yes

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19

	Top Five and Bottom Five Aptitudes	
APTITUDE:	**12. In your position, which aptitudes are the 5 <u>most</u> important? *(check 5 aptitudes)***	**13. In your position, which aptitudes are the 5 <u>least</u> important? *(check 5 aptitudes)***
a. Active learning		
b. Anomaly detection		
c. Category flexibility		
d. Complex problem-solving		
e. Convergent creative thinking		
f. Decision making		
g. Deductive reasoning		
h. Emotional intelligence		
i. Flexibility of closure		
j. Fluency of ideas		
k. Inductive reasoning		
l. Information and technology aptitude		
m. Information ordering		
n. Instructing		
o. Mathematical reasoning		
p. Memorization		
q. Mental agility		
r. Mental model ability		
s. Modeling program execution		
t. Number facility		
u. Oral Comprehension		
v. Oral Expression		
w. Originality		
x. Pattern recognition and scanning		
y. Perceptual Speed		
z. Problem Sensitivity		
aa. Psychomotor speed		
bb. Resistance to interference		
cc. Selective attention		
dd. Spatial visualization		
ee. Speed of closure		

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19

	Top Five and Bottom Five Aptitudes	
APTITUDE:	**12. In your position, which aptitudes are the 5 most important? *(check 5 aptitudes)***	**13. In your position, which aptitudes are the 5 least important? *(check 5 aptitudes)***
ff. Systems thinking		
gg. Time sharing		
hh. Time management		
ii. Vigilance		
jj. Visuospatial working memory		
kk. Written comprehension		
ll. Written expression		

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19

SECTION 3: TRAITS

Please review the list of traits (personal characteristics that can affect how well someone performs a job) provided below:

Aptitude	Definition
Achievement	The degree to which individuals establish and maintain personally challenging achievement goals and exerting effort toward mastering tasks.
Adaptability	The degree to which individuals are open to change (positive or negative) and to considerable variety in the workplace.
Analytical thinking	The degree to which individuals analyze information and use logic to address work-related issues and problems.
Attention to detail	The degree to which individuals are careful about detail and thorough in completing work tasks.
Concern for others	The degree to which individuals are sensitive to others' needs and feelings and being understanding and helpful on the job.
Cooperation	The degree to which individuals are pleasant with others on the job and display a good-natured, cooperative attitude.
Dependability	The degree to which individuals are reliable, responsible, and dependable, and fulfill obligations.
Independence	The degree to which individuals develop their own ways of doing things, guide themselves with little or no supervision, and depend on themselves to get things done.
Initiative	The degree to which individuals have willingness to take on responsibilities and challenges.
Innovation	The degree to which individuals use creativity and alternative thinking to develop new ideas for and answers to work-related problems.
Integrity	The degree to which individuals are honest and ethical.
Leadership	The degree to which individuals have a willingness to lead, take charge, and offer opinions and direction.
Need for cognition	The degree to which individuals enjoy participating in mentally demanding tasks.
Need for cognitive closure	The degree to which individuals have the need to arrive at a solution during problem solving
Persistence	The degree to which individuals have persistence in the face of obstacles.
Resilience	The degree to which individuals have the capacity to rebound or bounce back from adversity, conflict, and failure or even positive events, progress, and increased responsibility.
Self control	The degree to which individuals maintain composure, keep emotions in check, control anger, and avoid aggressive behavior, even in very difficult situations.
Self-discipline	The degree to which individuals tend be focused and dedicated to working hard and completing tasks in a timely manner.

Aptitude	Definition
Situational awareness	The degree to which individuals pay attention to their surroundings and rarely get lost or surprised.
Skepticism	The degree to which individuals distrust others.
Social orientation	The degree to which individuals prefer to work with others rather than alone and are personally connected with others on the job.
Stress tolerance	The degree to which individuals accept criticism and deal calmly and effectively with high stress situations.
Team player	The degree to which individuals prefer working in teams and help people work together better.
Tolerance for risk	The degree to which individuals are risk taking.

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19

14. Can you think of any traits that are important for your position that are missing from this list?

- ☐ **No**
- ☐ **Yes**

If you selected yes, please name the aptitude and briefly describe it.

Trait Name: ___________________________________

Description:

____
____
____
____
____
____
____
____
____
____
____
____
____

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19

PART 1 - Importance:

15. Using a scale from 1 to 5, how important is each trait to the performance in the position you indicated at the beginning of this survey. Importance refers to the significance of a trait for accomplishing one's job responsibilities.

PART 2 - Needed at entry:

16. Indicate whether is each trait is needed at entry (upon starting the position) or not needed at entry (can be learned or trained while in the position).

	PART 1: IMPORTANCE					**PART 2: NEEDED AT ENTRY**	
TRAIT:	**Not at all important**	**Slightly Important**	**Moderately Important**	**Very Important**	**Extremely Important**	**Not required at entry**	**Required at entry**
a. Achievement	1	2	3	4	5	No	Yes
b. Adaptability	1	2	3	4	5	No	Yes
c. Analytical thinking	1	2	3	4	5	No	Yes
d. Attention to detail	1	2	3	4	5	No	Yes
e. Concern for others	1	2	3	4	5	No	Yes
f. Cooperation	1	2	3	4	5	No	Yes
g. Dependability	1	2	3	4	5	No	Yes
h. Independence	1	2	3	4	5	No	Yes
i. Initiative	1	2	3	4	5	No	Yes
j. Innovation	1	2	3	4	5	No	Yes
k. Integrity	1	2	3	4	5	No	Yes
l. Leadership	1	2	3	4	5	No	Yes
m. Need for cognition	1	2	3	4	5	No	Yes
n. Need for cognitive closure	1	2	3	4	5	No	Yes
o. Persistence	1	2	3	4	5	No	Yes
p. Resilience	1	2	3	4	5	No	Yes
q. Self-control	1	2	3	4	5	No	Yes
r. Self-discipline	1	2	3	4	5	No	Yes

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19

TRAIT:	PART 1: IMPORTANCE Not at all important	Slightly Important	Moderately Important	Very Important	Extremely Important	PART 2: NEEDED AT ENTRY Not required at entry	Required at entry
s. Situational awareness	1	2	3	4	5	No	Yes
t. Skepticism	1	2	3	4	5	No	Yes
u. Social orientation	1	2	3	4	5	No	Yes
v. Stress tolerance	1	2	3	4	5	No	Yes
w. Team player	1	2	3	4	5	No	Yes
x. Tolerance for risk	1	2	3	4	5	No	Yes

TRAIT:	Top Five and Bottom Five Traits 17. In your position, which traits are the 5 <u>most</u> important? *(check 5 traits)*	18. In your position, which traits are the 5 <u>least</u> important? *(check 5 traits)*
a. Achievement		
b. Adaptability		
c. Analytical thinking		
d. Attention to detail		
e. Concern for others		
f. Cooperation		
g. Dependability		
h. Independence		
i. Initiative		
j. Innovation		
k. Integrity		
l. Leadership		
m. Need for cognition		
n. Need for cognitive closure		
o. Persistence		
p. Resilience		

	Top Five and Bottom Five Traits	
TRAIT:	**17. In your position, which traits are the 5 most important? *(check 5 traits)***	**18. In your position, which traits are the 5 least important? *(check 5 traits)***
q. Self-Control		
r. Self-discipline		
s. Situational awareness		
t. Skepticism		
u. Social orientation		
v. Stress tolerance		
w. Team player		
x. Tolerance for risk		

SYMBOLS, ABBREVIATIONS AND ACRONYMS

AAR – after-action review
ABM – Air Battle Manager
AC – Apprentice Course
ACE-R – Addenbrooke's Cognitive Evaluation
ACM – Association for Computing Machinery
A-DMC – Adult Decision-Making Competence
AF – Air Force
AFECD – Air Force Enlisted Classification Directory
AFRL – Air Force Research Laboratory
AFOCD – Air Force Officer Classification Directory
AFOQT – Air Force Officer Qualifying Test
AFPC – Air Force Personnel Center
AFSCs – Air Force Specialty Codes
AI – Aviation Information
AO – Assembling Objects
AR – Arithmetic Reasoning
ARI – Army Research Institute
AS – Automotive and Shop
ASA – advanced situational awareness
ASVAB – Armed Services Vocational Aptitude Battery
ATL – Adaptive Thinking and Leadership
BC – Block Counting
CATA – Cyber Aptitude and Talent Assessment
CASL – Center for Advanced Study of Language
CCTST – California Critical Thinking Skills Test
CE – circular economy
CIP – Career Intermission Program
CSF2 – Comprehensive Solider and Family Fitness
CSO – Combat Systems Officer
CT – Cyber Test
CT – Critical Thinking
DA – Deliberative Action
D-CITE – Decisions based on Collaborative Interactions in TEams
DHS – Department of Homeland Security
DoD – Department of Defense
DTIC – Defense Technical Information Center
ECD – evidence–centered design
EDPT – Electronic Data Processing Test
EI – emotional intelligence
ETA – Employment and Training Administration
ETS – Educational Testing Service
FA – Figure Analogies
GBAs – game–based assessments
GCHQ – Government Communications Headquarters

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19

GRE – Graduate Record Examination
GS – General Science
HR – heart rate
HumRRO – Human Resources Research Organization
IC – Instrument Comprehension
IE – Intellectual Efficiency
IEEE – Institute of Electrical and Electronics Engineers
IJSG – International Journal of Serious Games
InGEAR – Intelligent Game-Based Evaluation and Review
IS – information systems
JCAC – Joint Cyber Analysis Course
KAIT – Kaufman Adolescent and Adult Intelligence Test
KSAO – knowledge, skills, abilities, and other
KSAs – knowledge, skills, and abilities
MAGE – mechanical, administrative, general, electrical
MATSS – Managerial Assessment and Training Simulation System
MC – Mechanical Comprehension
MK – Mathematics Knowledge
MoCA – Montreal Cognitive Assessment
MSCEIT – Mayer-Salvoey-Caruso Emotional Intelligence Test
MTMM – multi-trait multi-method
MTT – Multitasking Test
NP – Newton's Playground
NPC – non-player characters
NS – Number Series
OAR – Occupational Analysis Report
O*NET – Occupational Information Network
PC – Paragraph Comprehension
PMESII-PT – political, military, economic, social, information, infrastructure, physical environment, and time
PS – Physical Science
PT – Proactive Thinking
QIS – quality of integration strategies
RC – Reading Comprehension
RT – Reactive Thinking
RTA – Real-Time Action
SA – situational awareness
SDI – Self-Description Inventory
SJT – situational judgment test
SMES – subject matter experts
STS – Systems Thinking Scale
STSR – Systems Thinking Scale Revised
TAPAS – Tailored Adaptive Personality Assessment System
TOEFL – Test of English as a Foreign Language
TR – Table Reading
UMD CASL –University of Maryland Center for Advanced Study of Language

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19

USDOL – US Department of Labor
USAF – United States Air Force
VA – Verbal Analogies
VBS3 – Training Virtual Battlespace 3 simulation
VE – Verbal Expression
VR – virtual reality
WCST – Wisconsin Card Sorting Task
WK – Word Knowledge

Distribution A: Approved for public release. 88ABW-2019-0557. Cleared 08/28/19