Technical Report 1422

Innovative Cognitive Assessment for Officer Classification

Michelle Wisecarver
Cory Adis
Personnel Decisions Research Institutes

Melissa Glorioso
U.S. Army Research Institute

Chelsey Byrd
Lia Engelsted
Jaclyn Martin
Personnel Decisions Research Institutes

January 2023

United States Army Research Institute for the Behavioral and Social Sciences

Approved for public release; distribution is unlimited.

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

Department of the Army
Deputy Chief of Staff, G1

Authorized and approved:

MICHELLE L. ZBYLUT, Ph.D.
Director

Research accomplished under contract
for the Department of the Army by

Personnel Decisions Research Institutes

Technical review by

Cristina Kirkendall, U.S. Army Research Institute
Jessica Carre, U.S. Army Research Institute

DISPOSITION

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

REPORT DOCUMENTATION PAGE

1. REPORT DATE (Month Year)	2. REPORT TYPE	3. DATES COVERED (Month Year)	
January 2023	Final	START DATE August 2020	END DATE August 2021

4. TITLE AND SUBTITLE
Innovative Cognitive Assessment for Officer Classification

5a. CONTRACT NUMBER	5b. GRANT NUMBER	5c. COOPERATIVE AGREEMENT NUMBER
W911NF-20-C-0047		

5d. PROGRAM ELEMENT NUMBER	5e. PROJECT NUMBER	5f. TASK NUMBER	5g. WORK UNIT NUMBER
622785	A790	1030	

6. AUTHOR(S)
Michelle Wisecarver, Cory Adis, Melissa Glorioso, Chelsey Byrd, Lia Engelsted, Jaclyn Martin

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

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)	11. SPONSOR/MONITOR'S REPORT NUMBER(S)
U.S. Army Research Institute for the Behavioral and Social Sciences 6000 6th Street (Bldg. 1464 / Mail Stop: 5610) Fort Belvoir, Virginia 22060-5610	ARI	Technical Report 1422

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

13. SUPPLEMENTARY NOTES
ARI Research POC: Dr, Melissa Glorioso, Selection and Assignment Research Unit

14. ABSTRACT
A literature review was conducted, for the U.S. Army Research Institute, on research that was relevant to developing assessment tools for 10 of the cognitive ability constructs identified in previous research as important for Army officers. Specifically, existing construct definitions and measures were examined, and recommendations were made regarding the most promising measurement approaches given a set of identified parameters. Recommendations are discussed for the measurement of each of the 10 constructs, which included combining some of the constructs with overlapping operational definitions. Proof-of-concept items for six constructs are presented.

15. SUBJECT TERMS
Officer cognitive ability, officer assessment, officer classification, leadership, analytical thinking, perceptive, systems thinking, problem solving, detail focused, technologically adept, interdisciplinary, multi-tasking, processes information and data

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES
a. REPORT	b. ABSTRACT	C. THIS PAGE		
Unclassified	Unclassified	Unclassified	Unclassified Unlimited	97

19a. NAME OF RESPONSIBLE PERSON	19b. PHONE NUMBER (Include area code)
Dr. Tonia Heffner	(703) 545-4408

Technical Report 1422

Innovative Cognitive Assessment for Officer Classification

Michelle Wisecarver
Cory Adis
Personnel Decisions Research Institutes

Melissa Glorioso
U.S. Army Research Institute

Chelsey Byrd
Lia Engelsted
Jaclyn Martin
Personnel Decisions Research Institutes

Selection and Assignment Research Unit
Tonia S. Heffner, Chief

January 2023

Approved for public release; distribution is unlimited.

ACKNOWLEDGMENT

We would like to thank Michelle Kaplan and Dr. Kimberly Wingert for their work in identifying information related to various analytical constructs and assisting with the conceptual thinking regarding the assessment of several of the constructs. We would also like to thank the members of the subject matter expert review panel for their advice and recommendations, and Cristina Kirkendall and Jessica Carre for their suggestions and improvements on an earlier version of this report.

INNOVATIVE COGNITIVE ASSESSMENT FOR OFFICER CLASSIFICATION

EXECUTIVE SUMMARY

Research Requirement:

The Army has identified cognitive ability constructs as important for officer success, but tools are not readily available to assess many of these constructs. Assessing these cognitive ability constructs has the potential to provide valuable feedback to Soldiers regarding these important capabilities and to provide valuable information to the Army regarding the classification of new officers into the various branches. A literature review was needed to collect information regarding how to assess 10 cognitive ability constructs that were identified in previous research as important for officer performance. These measures were identified as: Analytic Thinking, Analyze Data or Information, Detail Focused and Precise, Processes Information and Data, Problem Solver, Perceptive, Multitasking, Systems Thinking, Interdisciplinary, and Technologically Adept.

Approach:

A search of the literature was conducted for each of the 10 cognitive ability constructs identified as important for officer performance. The project team engaged in an iterative process of identifying relevant research then conducting collaborative reviews to determine where to focus further inquiry. An initial set of possible assessment approaches for each construct was proposed. A senior level subject matter expert (SME) review was conducted in which a panel of three researchers with Ph.D.s in Industrial-Organizational Psychology, each with 20 or more years of experience with assessment in the military reviewed the identified assessment approaches, provided feedback regarding the approaches, and suggested next steps. Based on the SME feedback, final assessment approaches were selected, and proof-of-concept items were developed.

Findings:

Operational measures were investigated that relate to the 10 cognitive ability constructs identified as important for officer success in the Army. In some cases, constructs were identified in the literature that were conceptually well matched to the defined Army construct, and in other cases, the constructs that appeared in the literature were only marginally related to the Army construct as it was defined.

One of our operating assumptions at the outset was that changes to the construct names and definitions should be avoided or minimized; in the end, however, recommendations were made to change the name and/or definition of six of the constructs for a variety of reasons. Analytic Thinking and Analyze Data or Information were combined to form a single construct labeled Analytic Thinking. Detail Focused and Precise and Processes Information and Data were also combined to form a single construct and labeled Detailed Information Processing. The definitions of Problem Solver and Multitasking were modified to be more appropriate for the construct that needed to be measured. Proof-of-concept items were developed for six constructs: Analytical Thinking, Detailed Information Processing, Problem Solver, Perceptive, Multitasking,

and Systems Thinking. It was recommended that two constructs, Interdisciplinary and Technologically Adept, not be developed at this time due to conceptual challenges with the construct definition and/or measurement challenges that arose given the constraints of the project assumptions or parameters.

Utilization and Dissemination of Findings:

Results from this research can be used to determine cognitive assessments that should be prioritized for development. Each of the six assessment concepts that were proposed would potentially be valuable elements of a cognitive battery for officers. While all six are recommended for further development if feasible, guidelines are provided that could be used to prioritize constructs for development. Once developed, these cognitive assessments can be used to select and classify officers to appropriate branches, as well as to provide feedback to officers for developmental purposes.

INNOVATIVE COGNITIVE ASSESSMENT FOR OFFICER CLASSIFICATION

CONTENTS

Page

INTRODUCTION 1

RESEARCH APPROACH 5

Assumptions 5
Research Strategy 6

CONSTRUCT REVIEW 6

Overview 6
Analytical Thinking 7
Analyze Data or Information 10
Processes Information and Data 12
Detail Focused and Precise 15
Multitasking 18
Problem Solver 21
Systems Thinking 26
Perceptive 30
Interdisciplinary 32
Technologically Adept 36

DISCUSSION 39

Analytical Thinking/Analyze Data or Information 39
Detailed Information Processing 40
Multitasking 40
Problem Solver 41
Systems Thinking 41
Perceptive 42
Interdisciplinary and Technologically Adept 43
Priorities for Development 43
Skill-based vs Self-report Measures 44
Conclusions 44

REFERENCES 46

APPENDCIES

APPENDIX A ANALYTICAL THINKING A-1

APPENDIX B DETAILED INFORMATION PROCESSING B-1

APPENDIX C MULTITASKING C-1

Page

APPENDIX D PROBLEM SOLVER D-1

APPENDIX E SYSTEMS THINKING E-1

APPENDIX F PERCEPTIVE F-1

LIST OF TABLES

TABLE 1. SEVEN COGNITIVE ABILITY CONSTRUCTS FROM THE USMA TALENT ASSESSMENT BATTERY 1

TABLE 2. TWENTY-TWO COGNITIVE ABILITY CONSTRUCTS AND DEFINITIONS IDENTIFIED BY ARI 2

TABLE 3. INITIAL 10 CONSTRUCTS AND THEIR DEFINITIONS 5

TABLE 4. DETAIL FOCUSED DEFINITIONS FROM THE LITERATURE 16

TABLE 5. ARMY CONSTRUCTS RELATED TO MULTITASKING 18

TABLE 6. MULTITASKING DEFINITIONS FROM THE LITERATURE 19

TABLE 7. ARMY TALENT ATTRIBUTE FRAMEWORK CONSTRUCTS RELATED TO PROBLEM SOLVING 22

TABLE 8. INTERDISCIPLINARY CONSTRUCT DEFINITIONS FROM THE LITERATURE 33

INNOVATIVE COGNITIVE ASSESSMENT FOR OFFICER CLASSIFICATION

Introduction

Success as a leader in the Army requires attributes that reflect Character, Presence, and Intellect (U.S. Department of the Army, 2019). Intellect includes key cognitive ability constructs such as mental agility, judgment, and innovation, which are required for success in making decisions, engaging in mission command, and working and adapting within the dynamic military environment. The Army has engaged in a number of efforts to identify and delineate the various cognitive and other skills and abilities required by Army officers for success. For example, the United States Military Academy identified 21 "Talents," or attributes that are important for students at the academy as they complete their degree and move to their first Army assignment (Boyle et. al, 2022). These 21 "Talents" include a variety of physical, social, motivational, and cognitive constructs and are assessed in the USMA Talent Assessment Battery (TAB). Cognitively-oriented skills and abilities in the TAB include Detail-focused, Interdisciplinary, Logical/Analytical, Multi-Tasker, Perceptive, Problem-solver, and Technologically Adept. Definitions for each of these seven constructs can be seen in Table 1.

Table 1

Seven Cognitive Ability Constructs from the USMA Talent Assessment Battery

Talent	Definition
Detail-Focused	Thorough, perceptive, and precise in all matters; Possesses a keen eye and notices everything
Interdisciplinary	Synthesizes and applies knowledge from multiple disciplines into a coherent overarching perspective
Logical/Analytical	Uses reason and thinks in terms of cause and effect; Able to deconstruct and solve complex problems
Multi-Tasker	Rapidly processes and prioritizes multiple demands simultaneously; Takes appropriate action when multiple things compete for his or her attention
Perceptive	Effectively discerns the deeper meaning or significance of one's observations (e.g., events, people, and communication)
Problem Solver	Able to choose between best practices and unorthodox approaches to reach a solution; Accomplishes the task
Technologically Adept	Understands and effectively uses the latest technologies

In another effort, the U.S. Army Research Institute for the Behavioral and Social Sciences (ARI) identified 63 skills and abilities that were potentially important for leader performance, including 22 identified in the cognitive domain, which can be seen in Table 2 (U.S. Army Research Institute, 2019).

Table 2

Twenty-two Cognitive Ability Constructs and Definitions Identified by ARI

Construct	Definition
Analytical Thinking	Analyzes information and uses logic to address work-related issues and problems
Analyze Data or Information	Identifies underlying principles, reasons, or facts by breaking down information or data into separate parts
Assessing and Mitigating Harm	Determines the potential for damaging outcomes and takes steps to minimize the risk of harm to self, others, or equipment (e.g., performing cost-benefits analysis; mandating training to mitigate risk, etc.)
Attentiveness	Focuses on the problem or situation and shifts attention between activities when appropriate
Awareness of Cognitive Biases	Realizes one's personal biases, predispositions, or inclinations in thinking when looking for solutions to problems
Creative Problem Solving	Develops and utilizes new and creative methods or strategies to accomplish work or achieve goals when established methods and procedures are inapplicable or ineffective
Focus	Mentally processes multiple sources of sensory information/data at the same time while avoiding distractions (e.g., flying a helicopter or commanding an armored vehicle)
Interdisciplinary Reasoning	Understands and integrates information from multiple professional disciplines to complete tasks and projects
Judgement and Decision Making	Makes decisions based on accurate and appropriate assessment of the costs/benefits and short- and long-term consequences of alternative actions and solutions; Makes timely decisions with incomplete information, while refraining from making hasty decisions in the absence of necessary information
Pattern Recognition	Perceives things quickly and accurately, and detects similarities or differences in objects, words or numbers
Precision	Attentive to detail and thorough, accurate, and precise in completing a task
Processes Information and Data	Compiles, codes, categorizes, calculates, tabulates, audits, or verifies information or data

Table 2, *(continued)*

Construct	Definition
Quantitative Reasoning	Solves problems that involve mathematical concepts or numbers
Reflective Thinking	Regularly and actively reflects on the connections between new information and existing knowledge
Situational Awareness	Perceives what is happening in the immediate environment, and how information, events, and actions will impact current and near-term goals and objectives
Spatial Orientation	Knows one's physical location in relation to the environment or knows where other objects are in relation to oneself
Spatial Visualization	Mentally visualizes objects and imagines how they would be positioned if moved or rotated in different ways
Strategic Thinking	Develops a complex, systems-level understanding of the relationship between his/her Army unit or organization and the broader environment and uses that understanding to envision a desirable future state for the unit/organization
Structured Problem Solving	Analyzes readily obtained information and evaluates results to select the best solution from a set of existing approaches to solve a problem
Systems Thinking	Conceptualizes and understands relationships and arrangements within and between relevant components and structures
Unstructured Problem Solving	Identifies complex problems, gathers related information, evaluates information relevance, evaluates the credibility of alternative information sources, and generates alternative solutions
Verbal Reasoning	Reasons and draws conclusions based on verbal or written materials

ARI conducted preliminary research to determine the importance of both the 21 "Talents" and 63 skills and abilities for officers in their jobs (U.S. Army Research Institute, 2019). Lieutenants assigned to 18 different Army branches rated the importance of each attribute for successful performance of their current job, using a scale that ranged from "0" (*Not Important at All*) to "4" (*Extremely Important*). Mean responses for four of the cognitive attributes from the TAB (Detail Focused, Logical/Analytical, Multi-tasker, and Problem Solver) were between "3" (*Very Important*) and "4" (*Extremely Important*) for all or nearly all branches. The other three cognitive attributes - Interdisciplinary, Perceptive, and Technologically Adept - generally had mean ratings between a "2" (*Moderately Important*) and "3" (*Very Important*), with one or two branches showing mean ratings that were higher, between "3" and "4."

Results were similar for ratings of the 22 cognitive skills and abilities identified by ARI, with lieutenants rating nine of the cognitive attributes (Focus, Attentiveness, Precision,

Situational Awareness, Verbal Reasoning, Analytical Thinking, Structured Problem Solving, Unstructured Problem Solving, and Judgement and Decision Making) between "3" (*Very Important*) and "4" (*Extremely important*) for all or nearly all branches. Eight of the cognitive attributes had mean ratings between a "2" (*Moderately Important*) and "3" (*Very Important*) for most branches, with a few branches showing mean ratings that were higher, between "3" and "4." Ratings for three attributes (Process Information and Data, Analyze Data or Information, Systems Thinking) were evenly split, with lieutenants from about half of the branches rating the attribute between and "2" and "3," and the other half rating the attribute between a "3" and "4."

These data emphasize that cognitive abilities are generally important for leaders across all Army branches; however, the data also suggest that certain types of cognitive abilities may be more important for some branches than others. As an example, Processing Information and Data from the 22 ARI cognitive attributes was rated as significantly more important by lieutenants in Field Artillery ($M = 3.27$) than those in Infantry ($M = 2.69$), $t(187)=4.18$, $p < .05$, and Systems Thinking was rated as significantly more important by lieutenants in Air Defense ($M = 3.17$) than Military Police ($M = 2.5$), $t(68)=2.90$, $p < .05$ (U.S. Army Research Institute, 2019). This suggests that in addition to using cognitive attributes to identify individuals who will succeed as Army officers, some of the attributes could potentially be useful in determining which officers would be better suited for success in specific branches.

Although the Army has identified these cognitive constructs as important for officer success, tools are not readily available to assess officers on many of these constructs. Improving the assessment of officers across these cognitive attributes has the potential to provide useful developmental feedback to officers regarding these important capabilities, and to provide valuable information to the Army regarding the classification of new officers into the various branches. In order to gain a better understanding of these important cognitive attributes, ARI has begun to conduct research to identify potential ways to measure them in officers. Toward this end, ARI identified 10 constructs from the above described 29 constructs for further examination as part of this research effort. Using existing ARI job analysis data, these constructs were selected for their relevance to officer performance, as well as their potential ability to differentiate performance in different branches.

The 10 constructs and their definitions based on the Army Talent Attribute Framework (ATAF) can be seen in Table 3 (Royston et. al, in preparation). The ATAF framework was developed through an initiative that engaged members of ARI as well as other Army stakeholders and subject matter experts in a process of compiling, classifying, and condensing attributes identified as important for Army leader success into a cohesive framework (Goodwin et al, 2020). This was an iterative process and definitions were taken from the ATAF available at the time (Royston et. al, in preparation), so may differ slightly from the final form of the ATAF once it has been completed. The objective of this report is to describe existing measures of these 10 constructs and provide recommendations for development and measurement strategies for Army owned assessments of these constructs.

Table 3

Initial 10 Constructs and Their Definitions

Construct	Definition
Analytical Thinking	Analyzes information and applies general rules and logic to address work-related issues and problems
Analyze Data or Information	Identifies underlying principles, relationships, general rules, reasons, or facts by breaking down information or data into separate parts
Detail Focused and Precise	Attentive to detail and thorough, accurate, and precise in completing a task
Processes Information and Data	Compiles, codes, categorizes, calculates, tabulates, audits, or verifies information or data
Problem Solver	Capacity to choose between best practices and unorthodox approaches to reach a solution; Accomplishes the task
Perceptive	Effectively discerns the deeper meaning or significance of one's observations (e.g., events, people, and communication)
Multi-Tasking	Rapidly processes and prioritizes multiple demands simultaneously; Takes appropriate action when multiple things compete for his or her attention
Interdisciplinary	Synthesizes and applies knowledge from multiple disciplines into a coherent, overarching perspective
Systems Thinking	Conceptualizes and understands relationships and arrangements within and between relevant components and structures
Technologically Adept	Learns how to use and apply advances in technologies or technological systems

Research Approach

Assumptions

While examining assessment approaches for each of these 10 constructs, there were a number of assumptions that guided our research approach. First and foremost was that the objective of our research was a targeted review that would focus on identifying existing methods to measure the 10 constructs, toward the end of making recommendations for developing an assessment for the Army. A second assumption was that, while changes to the construct name and definition would likely occur, these changes should be avoided or minimized when possible. Third, recommendations regarding measurement strategies should focus on measuring cognitive skills or abilities directly rather than self-reported behaviors. Fourth, the assessments should not require any specialized or Army-specific knowledge. Finally, recommendations regarding measurement strategies should minimize the time requirements for test takers to the extent possible, and although the assessments are expected to be administered using a computer-based

platform, the requirements and complexity of the systems needed to host and take the assessment should be minimized.

Research Strategy

For each of the 10 constructs an initial search of the literature was conducted using relevant key terms to obtain a sense of the breadth and depth of the available research. In some cases, the available literature was too broad and had to be narrowed, and in other cases the literature was lacking or too narrow and needed to be broadened using additional terms in order to facilitate the search for relevant approaches. The project team engaged in an iterative process of identifying relevant research and then conducting collaborative reviews to determine the most relevant areas of research on which to focus. In some cases, decisions needed to be made regarding aspects of the construct definition or types of assessments on which to focus. An initial set of possible assessment approaches for each construct was then proposed, with corresponding advantages and disadvantages. A senior level subject matter expert (SME) review was conducted in which a panel of three researchers with Ph.D.s in Industrial-Organizational Psychology, each with 20 or more years of experience with assessment in the military reviewed the identified approaches and provided feedback regarding the assessment approaches and suggestions for next steps. Although each of the 10 constructs were conceptually unique, in some cases it was unlikely that the modest conceptual differences could be reliably distinguished in an operational measure, and recommendations were made to combine constructs for the purpose of measurement. Based on the SME feedback, final assessment approaches were selected, and proof-of-concept items were developed. In the following section the research, decisions, and measurement recommendations related to each construct will be described.

Construct Review

Overview

In this section the research and decisions related to each of the 10 constructs are described. For each construct we first describe the initial research findings and decisions that were made regarding aspects on which to focus, as well as related or overlapping constructs. Next, the initial assessment approaches that were considered for the construct are provided as well as any decisions and guidance offered by the senior SME review panel. Finally, any modifications made to the construct definition are described, recommendations are provided for the measurement of the construct, and proof-of-concept items for six of the constructs are presented in Appendices A through F.

Analytical Thinking

Analytical Thinking refers to the tendency to use analytical skills to address problems. Specifically, this construct was defined as “analyzes information and applies general rules and logic to address work-related issues and problems” (Royston et al., in preparation). Analytical Thinking can be conceptualized as the thinking and reasoning processes to solve problems and make decisions (Gantt & Natt, 2006). Initial searches indicated that a unified framework does not currently exist in the literature for the conceptualization or measurement of Analytical Thinking. Related constructs identified using various terms associated with analysis and logic included:

business intelligence analytics (e.g., Corte et al., 2014), analytical skills (e.g., Chiang et al., 2012), and logical thinking/reasoning (e.g., Engelmann et al., 2020; Prahalad & Bettis, 1986; Wilhelm, 2005). Associated measures ranged from (a) personality-type measures focused on thinking styles to (b) skill-based measures focused on analytical skills such as statistical analysis and general business intelligence to (c) cognitive-based measures focused on assessing different types of logical reasoning (inductive, deductive, abductive).

Personality-type operationalizations regarding assessing thinking style to evaluate the extent to which a person thinks analytically versus intuitively; conceptualized similarly to a personality trait. For instance, Epstein and colleagues (1996) measured "analytical-rational thinking" (i.e., a deliberative, systematic, and logical form of thinking) using items from a Need for Cognition scale, such as, "I would prefer complex to simple problems." Allinson and Hayes (1996) developed a 38-item Cognitive Style Index designed to measure where those in managerial and professional occupations fall on the dimension of analytical versus intuitive thinking.

The skill-based measures and conceptualizations related to Analytical Thinking typically centered around business intelligence and statistical analysis (e.g., Chiang et al., 2012; Corte et al, 2015; de Queiroz Tourinho et al., 2019). For example, one skill-based operationalization is the ACOMP (analytical competence) scale which includes items reflecting analytical skill, information technology (IT) and data skill, and problem-solving skill (de Queiroz Tourinho et al., 2019).

Cognitive ability-based measures focused on various forms of reasoning, including measures of deductive, inductive, and abductive reasoning. The two most commonly researched forms of reasoning are defined as follows: deductive reasoning - "the ability to apply general rules to specific problems to produce answers that make sense," and inductive reasoning - "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.). Deductive reasoning is an established construct with several measurement approaches. For example, it is often measured using logical syllogism items that ask test takers to arrive at conclusions based on two or more propositions or rules (e.g., The Syllogistic Test, Kuhn, 1977; Syllogism Test, Gottesman & Chapman, 1960). In another variation on syllogistic reasoning, Johnson-Laird's (1995) reasoning task provides test takers with 23 syllogisms and 23 implications, disjunctions, and conjunctions. Test takers determine if a third sentence is appropriate based on the first two sentences. Inductive reasoning is often measured using a serial completion task (e.g., Girelli, Semenza, & Delazer, 2004; Holzman, Pellegrino, & Glaser, 1983; Kotovsky & Simon, 1973; Lefevre & Bisanz, 1986; Simon & Kotovsky, 1963) where a series of letters or numbers are presented, and the test taker must indicate what letter or number occurs next in the pattern.

The third, and less well known, type of reasoning is called abductive reasoning and it relates to a person's ability to develop and evaluate logical inferences or hypotheses in the absence of complete information. While both deductive and inductive reasoning show how conclusions are formed using evidence in different ways (derived vs. inferred, respectively), abductive reasoning more closely parallels the reasoning process that individuals conduct on a daily basis (Colberg et al., 1985; Fischer, 2001; Mellouli & Bouchon-Meunier, 2003). In abductive reasoning, conclusions must be qualified as best or "optimal" (rather than correct or

incorrect) under conditions of doubt and uncertainty, which is commonplace in regular decision-making as real-world environments generally lack absolutes.

Based on discussions within the project team, the reasoning measures provide the best fit for a potential measurement of Analytical Thinking, given that they focus on measuring the construct directly rather than through self-reported styles or approaches. In this way, the reasoning measures enable assessment of concepts not just numeric data skills and minimize overlap with other cognitive constructs such as problem solving. Because ARI already has measures of inductive and deductive reasoning, the decision was made to focus on measures of abductive reasoning. In addition, from a practical standpoint, this focus is useful because abductive reasoning is the type of reasoning that officers will use on a daily basis to solve work-related issues and problems.

Further research was then conducted to collect additional information specifically regarding abductive reasoning. Research regarding how to measure abductive reasoning spans disciplines from philosophy to psychology to computer science (e.g., Fischer, 2001; Mellouli & Bouchon-Meunier, 2003; Schvaneveldt & Cohen, 2010; Wilhelm, 2005). Some measures assess individual reactions to future scenarios (Fisher, 2001; Mellouli & Bouchon-Meunier, 2003) and others include computational tools that measure hypothesis development (Schvaneveldt & Cohen, 2010).

Recent research has focused on the development of abductive reasoning measures that capture two phases of the reasoning process: hypothesis generation and hypothesis evaluation (Kwon et al., 2009; Lee, 2012; Zelechowska et al., 2020). To capture authentic, unrestricted hypothesis generation, test taker input must be collected in a constructed response (i.e., open-ended, short answer) format. This format would either need to be scored by hand or would require specialized essay scoring software and would be expensive and time consuming to develop and administer. It is also possible that writing ability could confound test-takers' scores under this operationalization.

While some assessments of abductive reasoning have used extensive constructed response measures to capture hypothesis formulation and testing (e.g., Zelechowska et al., 2020), other approaches have used formats that are more easily scored. Kingry et al. (2020), for example, used a multiple-choice format in which the item stems provided an incomplete set of observations or a short scenario. Test takers are instructed to select the best option among response options that list possible explanations, or hypotheses about the described observations/scenarios. Further depth of measurement is achieved by asking respondents to identify one question out of a group of questions whose answer would yield the most helpful information to understand what is happening in the scenario. This provides a more comprehensive assessment of test taker reasoning by examining both the process the test taker is using to evaluate the hypotheses (i.e., what questions are helpful) in addition to examining their reasoning outcome (i.e., the best explanation). Thus, the constructed response formats can measure both hypothesis formulation and hypothesis evaluation but require hand scoring or specialized essay scoring programs; whereas the multiple-choice approach measures only hypothesis evaluation but can be scored easily and automatically.

Measurement Strategy Recommendations

Multiple assessment formats were described to the SME review panel as possible measurement approaches for abductive reasoning, including a computer-scored short answer, a situational judgement test (SJT) format, and the multiple-choice hypothesis evaluation measure described in Kingry et al. (2020). The SME review panel, composed of Industrial and Organizational P.h.D.s from ARI and PDRI, agreed with the usefulness of focusing on abductive reasoning to measure Analytical Thinking and recommended moving forward with the multiple-choice hypothesis evaluation approach. In addition, due to conceptual overlap between Analytical Thinking and other cognitive constructs, the recommendation was made to consider whether a measure of Analytical Thinking could be successfully combined with a measure of any of the other nine constructs, thus minimizing test time and maximizing test efficiency.

The project team identified the construct Analyze Data or Information as sufficiently overlapping with the Analytical Thinking construct to enable the development of a single measure that would capture information related to both constructs. Specifically, a single measure is proposed that focuses on abductive reasoning, but also includes items related to evaluating information, a theme identified for the Analyze Data or Information construct. In all, four item types were proposed to measure the three subdimensions of the overlapping construct: information evaluation, hypothesis evaluation, and process evaluation. The first three item types would use a common scenario or incomplete set of observations and then asking a series of three questions about it. The fourth item type is measured independently from the first three.

For the first item type, respondents are provided with an item stem that includes an incomplete set of observations or a short scenario. Respondents are asked the extent to which the information provided is sufficient to determine an explanation for what happened in the scenario. This item type is designed to measure the information evaluation subdimension by assessing how well a person evaluates the sufficiency of the information at hand. Next, in the second item type, respondents are instructed to select the most plausible explanation for the scenario from response options that list possible explanations, or hypotheses about the scenario. Again, the scenario is the same one presented for the first item. The second item type is designed to measure the hypothesis evaluation subdimension by assessing how well a person evaluates the plausibility of different hypotheses. In order to get more data points out of each scenario, additional modifications could be presented to the respondent at this point, and the respondent could be asked to again select the most plausible explanation based on the modification presented to them. In the third item type, the item stem asks which of the questions test takers would pose first if they wanted to better understand what happened in the given scenario. This item type is designed to measure the process evaluation subdimension by assessing how well a person gathers additional information to refine their hypothesis.

The final item type is independent from the first three. In the fourth item type, the stem provides a statement that takes a stance on a certain issue and test takers identify the most and least effective counterpoints. This item type is designed to measure the information evaluation subdimension by assessing how well a person evaluates the sufficiency of the information at hand. Further details as well as example items can be seen in Appendix A.

There is not yet definitive information regarding the relative importance of the three subdimensions or the validity of the four different item types in capturing information about the three subdimensions. While these subdimensions and measurement approaches show promise conceptually by capturing information about an individual's abductive reasoning ability, empirical research will be required once an operational measure is developed to determine the relative contribution of the dimensions and usefulness of the various measurement approaches.

Analyze Data or Information

The definition provided for Analyze Data or Information was "identifies underlying principles, relationships, general rules, reasons, or facts by breaking down information or data into separate parts" (Royston et al., in preparation). As mentioned, there is a high degree of conceptual overlap between this construct and Analytical Thinking, which was defined as, "analyzes information and applies general rules and logic to address work-related issues and problems." Both definitions involve identifying general rules based on a set of information. Similar to Analytical Thinking, there is not a unified framework in the literature for the conceptualization of the ability to Analyze Data or Information. Related constructs include data analytics competency/skills (e.g., Ghasemaghaei et al., 2018; Jafar et al., 2016), information gathering (Brubacher et al., 2019), and judgment of information (Rieh & Belkin, 1998). Constructs related to analyzing data, such as those that assess data analytics skills and knowledge, are popular in the business, management, and information science literatures (e.g., Boyatzis & Kolb, 1995; Ghasemaghaei et al., 2018; Ghasemaghaei, 2019; Gil, 2016; Hattingh et al., 2019; Jafar et al., 2016; Russom, 2011; Tippins & Sohi, 2003; Wang & Strong, 1996).

Data analytics research considers how individuals combine processes and tools to engage in various analytical tasks (e.g., Ghasemaghaei et al., 2018; Ghasemaghaei, 2019; Wang et al., 1996). Measures evaluate individuals' skills with statistics and data mining as well as their knowledge and sophistication with varying tools (e.g., Ghasemaghaei et al., 2018). One area of data analytics research focuses on skill with big data analytics specifically, examining individuals' understanding and appropriate analysis of big data sets while also ensuring decision-making quality (e.g., Gil, 2016; Hattingh et al., 2019; Russom, 2011). Given that these concepts and measures tend to center around or involve mathematical skills and knowledge of specific tools, they were not well suited for our purposes.

The concept of information gathering and the judgment of information, such as understanding the quality and authority of information sources was more aligned with our objective of focusing on assessing how well respondents can analyze information, rather than analyzing quantitative data. Individuals must be able to judge the information they receive by appropriately examining and analyzing the quality of the information and whether the information is useful in a given situation utility (Rieh & Belkin, 1998; Wang & Strong, 1996). There are three main bodies of literature on information gathering. The first body of literature focused on specific physiological issues of perception and cognition (e.g., Lucas et al., 2019). The second body of literature focused on information gathering with respect to interrogation (e.g., Meissner et al., 2014; Miller et al., 2018). Finally, the third body of literature examined information gathering that is focused on gathering information in a given domain, such as applying to college (Meyer & Cranmore, 2020), using the internet (Tippins & Sohi, 2003), or conducting clinical interviews (Brubacher et al., 2019). From the perspective of measurement,

these measures often use qualitative interviews to measure information gathering rather than a standardized assessment tool. For example, Brubacher et al., 2019 used semi-structured interviews with clinicians to examine information gathering in the medical field, and Rieh and Belkin (1998) assessed facets of judgment of information quality such as the source, content, and accuracy using interviews regarding searches of the worldwide web. One measure that did not use an interview is the 72-item Q-Sort to identify information gathering and analysis dimensions (Boyatzis & Kolb, 1995). Although the Q-Sort supported the relevance of information gathering and information analysis to learning skills for management populations it did not provide a useful tool for directly assessing those skills in leaders.

None of the various bodies of research provided practical tools to assess the Analyze Data or Information construct, despite supporting the importance of understanding how individuals gather, analyze, and utilize information for differing purposes. The project team used the above concepts to develop ideas for more standardized measurement approaches.

Measurement Strategy Recommendations

In the absence of a clear direction from the literature regarding the measurement of the Analyze Data or Information construct, the project team focused initial recommendations for an assessment on the actions specifically described in the definition of the construct: identifying the underlying principles, relationships, general rules, reasons, or facts by breaking down information. Two potential options were identified that used a task-based approach in which test takers would receive information and would respond to questions about said information that would test their understanding of underlying principles, relationships, and rules in the information provided to them. One measurement proposal was for an assessment that includes a long prompt with multiple questions associated with the prompt. The prompt could be similar to one that intelligence report officers might receive that provides information about events and people in a fictitious setting. Test takers would respond to questions that tested their understanding of the underlying relationships, norms, etc. that were described in the report. This approach would directly measure the cognitive skills associated with this construct and may be viewed as face valid due to the realism of the task. Possible disadvantages, however, would be that the assessment could be lengthy as test-takers would require time to read through the background materials (e.g., intel report) and the nature of the long prompt would likely cause the measure to reflect reading comprehension in addition to information analysis, which could confound the measure. An alternative to the prompt with a single long intelligence report could be using a series of shorter tasks as prompts that require breaking down information into separate parts to answer prompts. For example, a task could be given to assign a list of people to training times that work, given certain scheduling and relational constraints. This would provide more independence from reading comprehension, increase the number of measurement opportunities, and allow scoring using Generalized Partial Credit Model (GPCM; Muraki, 1992, 1993), which has the potential to generate more item information and lead to lower standard error at the test level.

Members of the SME panel agreed with the potential appropriateness of the task-based assessment approach; however, they recommended prioritizing the importance of reduced administration time and exploring the possibility of capitalizing on the conceptual overlap of the analyze data and information construct with the Analytical Thinking construct and combining the

two constructs into a single measure. After reviewing the two construct definitions and measurement recommendations, the project team recommended maintaining the label "Analytical Thinking" for the combined construct and using the operational definition, "analyzes information and applies general rules and logic to evaluate the information at hand, which is often incomplete, and comes to the best possible conclusion or decision." A combined measurement was proposed to cover the two constructs that focuses on abductive reasoning but includes items that require judging the completeness of the available information. As described in the section on Analytical Thinking, five item types were proposed that relate to the three subdimensions - information evaluation, hypothesis evaluation, and process evaluation. Details about the item types and a set of example items are provided in the previous section and in Appendix A.

Processes Information and Data

The construct Processes Information and Data was defined as "the ability to compile, code, categorize, calculate, tabulate, audit, or verify information or data" (Royston et al., in preparation). The definition clearly lays out the dimensions or components making up the construct, so one potential direction for this measure would be to test performance on each of those components and create composite scores for the construct. Another approach would be to focus on the skills and/or cognitive processes that are common across those components. Such an approach would provide a more general view of the construct.

Information processing is ubiquitous throughout modern cognitive psychology (Lord & Maher, 1990; Lord, 1985). It is the computation that humans and computers perform enabling them to take in, transform, and use information. While some of the literature coming out of the field of cognitive psychology deals with individual differences in information processing, most is concerned with overall commonalities across individuals in how information processing occurs (e.g., most takes a nomothetic approach rather than idiographic approach). Beyond cognitive psychology, research on information processing spans a wide variety of disciplines including marketing communications (e.g., Kitchen et al., 2003; Sojka & Giese, 2006) management (e.g., Lord, 1985; Lord & Maher, 1990; Walsh, 1995; Ungson et al., 1981), and STEM education (e.g., Burns et al., 1985; Özgelen, 2012).

Walsh (1995) reviewed two different modes of information processing – theory-driven, or top-down processing, and data-driven, or bottom-up processing – which reflect two directions of information flow. In top-down processing, information from general concepts and ideas have an impact on lower-level cognition such as perception or attention. In bottom-up processing, information from perception and attention is assembled and arranged into higher-order cognition. Both forms of cognition are important and both modes are used to process information, however, for present purposes, a reasonable focus is on bottom-up processing. This focus is based on the definition provided by the Army Talent Management Framework (Royston et al., in preparation); the Processes Information and Data construct is conceptualized as data-driven information processing. This definition focuses on one's ability to deal with smaller bits of data without reference to larger theories or overarching goals; for example, compiling, coding, and categorizing information in order to maintain its form, rather than to reach conclusions.

Because of the ubiquity of information processing in cognitive psychology, the field of cognitive psychology does not have a single construct called processes information; rather, there are constructs in the literature that are conceptually related to processing information. For example, there are several constructs related to processing speed (Faust, Balota, Spieler, & Ferraro, 1999; Kyllonen, 1985; Lohman, 1989; Sheppard & Vernon, 2008). For example, response latency and reaction times have been studied for over a century, since Donders' method of subtraction (see Gottsdanker & Scragg, 1985 for a description of Donders' method from the late 1800s). More recently, other information processing skills such as visual search (Eriksen & Shultz, 1979) and inspection time (Chaiken, 1994) have been identified as stable individual differences.

Several constructs relate conceptually to the processing information and data construct. Analytic competence (de Queiroz Tourinho, Sanchez, & Brown, 2019) and Analytic Thinking (Gantt & Natt, 2006) both have overlap with processing information and data, however, these analytic concepts are captured elsewhere with other constructs. Data analytics competency (Ghasemaghaei, Ebrahimi, & Hassanein, 2018; Tippins & Sohi, 2003) is also similar to Processes Information and Data. With this competency, the focus is on both the components identified in the definition and processes that are more mathematically oriented such as predictive analytics, statistics, and data mining.

One measure of information processing that was developed in the education field is called the Test of Integrated Process Skills (TIPS), a multiple-choice assessment developed to assess science process skills, such as formulating hypotheses, identifying relevant variables, and interpreting data (Burns et al., 1985). The test was specifically designed for measuring these process skills within the integrated science context and was therefore designed to span multiple scientific disciplines. As such, test takers were provided with information from different disciplines and contexts (e.g., auto efficiency, marine biology, food production, agriculture). To the extent that a test taker already had knowledge about these science fields, however, they would likely score higher on the test. This issue creates the potential for test takers to have an unfair advantage if they possess background knowledge on a subject on which they are tested.

While these conceptualizations of information processing are relevant, the definition of Processes Information and Data focuses on a list of specific processing skills or abilities, rather than the broader process or steps of attending to, storing, retrieving, and making judgments about information (see the definition in Table 3). Specifically, the definition indicates that an individual should be able to compile, code, categorize, calculate, tabulate, audit, and verify information or data. In addition, the construct Analyze Data or Information already focuses on general aspects of processing information.

Given these factors, the project team recommended that the measure of the Processes Information and Data construct focus on measuring the specific skills listed in the definition: compile, code, categorize, categorize, tabulate, audit, and verify. Definitions were identified for each of these terms (Merriam-Webster, 2021):

- Compile: to compose out of materials from other documents
- Code: to put in or into the form or symbols of a code (code = a system of signals or symbols for communication)

- Categorize: to put into a category, classify (category = a division within a system of classification)
- Tabulate: to count, record, or list systematically
- Audit: a methodical examination and review
- Verify: to establish the truth, accuracy, or reality of the claim

In addition, the decision was made to avoid focusing on mathematical or statistical skills and to focus instead on the processing of information rather than the processing of data. Compiles, codes, categorizes, audits, and verifies information were the components deemed least mathematically oriented.

In order to capture the different components of the definition, the project team developed the concept for a task-based assessment approach in which test takers would be presented with a small task (e.g., figuring out parts of a graph, making entries into a spreadsheet) and would respond to questions that tested their ability to perform some of the skills from the construct definition (compile, code, categorize, tabulate, audit, verify). The benefits to this approach are that it would directly measure the various cognitive skills identified in the construct definition and that it could produce scores for specific subdimensions, if desired, though that would require a greater number of items and a longer test. This type of assessment would also allow for a range of item difficulties, enabling the use of a computer-adaptive assessment approach, if that was desired. In addition, the items could be scored using the GPCM (Muraki, 1992, 1993), which has the potential to generate more information from each item. A potential disadvantage to this approach is that there could be a possible confound with reading comprehension, depending on the nature of the information stimuli provided to test takers.

Measurement Strategy Recommendations

The SME review panel agreed with the usefulness of the task-based approach. Following a review of the constructs, sufficient task similarity was identified between the proposed measure of Processes Information and Data and the proposed measure of Detail Focused and Precise. Detail Focused and Precise is defined as, attentive to detail and thorough, accurate, and precise in completing a task (Royston et al., in preparation). Both constructs could be measured with a similar task-based approach, so a single measure was proposed for a combined construct that reflects detail-focused processing of information through a measure that incorporates tasks requiring processes such as: compiling, categorizing, auditing, and verifying.

The project team developed a concept to use military-type objects such as reports or simple graphics as stimuli in which test takers will answer questions that require identifying errors and problems (see Appendix B). For example, in the example scenario included in Appendix B, the test taker is prompted to complete a DA 638 form to recommend a Soldier for an award. The DA 638 form provides a context for information that officer candidates will be unfamiliar with at the time they take the assessment, but also represents a task they would realistically perform as a part of their job. Thus, this scenario should be face valid and engaging, without providing an unfair advantage based on past experience. Additionally, both the background information and the respective response options could be constrained to reduce reading requirements (e.g., restrict the number of awards available), which would reduce the time required for administration and reduce the potential for a reading comprehension confound.

The different item types in the proposed measure in Appendix B reflect detail focused and the Processed Information and Data component skills *compile, categorize, audit,* and *verify*. Additionally, a method for capturing the component skill *tabulate* in one of the items is provided in the test specification notes. Additional information in the following section provides a review of the literature on traits such as detail focused, precision-oriented, and related constructs.

Detail Focused and Precise

The construct Detail Focused and Precise is defined in the ATAF as "Attentive to detail and thorough, accurate, and precise in completing a task" (Royston et al., in preparation), however, across the various Army documents reviewed for this project differing construct names and definitions were found. The USMA list of 21 Talents included a construct called "Detail-focused," defined as "Thorough, perceptive, and precise in all matters; Possesses a keen eye and notices everything," and the 22 ARI cognitive constructs included one labeled, "Precision" that had the same definition as the ATAF construct, Detail Focused and Precise.

Literature examining topics such as "detail focused" and "attention to detail" span a number of psychological disciplines, including literature focused on autism (e.g., Happé & Frith, 2006; Barca et al., 2010), creativity (e.g., Kirton, 1976; Miron et al., 2004), cognitive/experimental psychology (e.g., Baylis, 1994; Baylis & Driver, 1992; Grimes, 1996), and quality control (e.g., Alefari et al., 2018). Each focuses on slightly different aspects or interpretations of the construct. Related constructs and their definitions that were identified in the literature can be seen in Table 4.

Though not directly relevant to developing a measure of detail focus for the Army, research focused on autism and creativity has suggested that there may be a curvilinear relationship between detail focus and performance such that focusing too intently on details can lead to a decrement in performance. Research examining detail focus in people on the autism spectrum indicates that they may be unable to process a global understanding of a situation or object because they are so intensely focused on small details; for example, they are able to determine the pitch of the sound of a cork being popped, but this distracts them from hearing or understanding a conversation (Happé & Frith, 2006; Barca et al., 2010). Similarly, research on creativity suggests that if an individual is too focused on details, their creative performance will be lower (Kirton, 1976; Miron et al., 2004; Miron-Spektor et al., 2011; Naveh & Erez, 2004; O'Reilly et al., 1991; Sok & O'Cass, 2015). Thus, while many cognitive constructs would demonstrate a purely linear relationship with officer performance, detail focus may be a construct for which there is a limit to its positive effects.

Table 4

Detail Focused Definitions from the Literature

Construct Name	Definition	Source Literature
Detail Focused Processing Bias	Inability to experience wholes without full attention to the constituent parts	Autism (Happé & Frith, 2006)
Attention to Detail	A cognitive style commonly associated with autism in which one attends more to fine-grained details of the world at the expense of more integrative, gestalt perceptions	Autism (Happé & Frith, 2006)
Attention to Detail	Performs tasks precisely, systematically, and reliably and thoroughly solves problems; Is thorough, masters all details painstakingly, enjoys detailed work	Creativity/ Innovation (Sok & O'Cass, 2015)
Detail Focus	Precision and accuracy; one of the organizational values that comprise the organizational culture	Creativity/ Innovation (Naveh & Erez, 2004)
Change Blindness/Change Detection	When changes to a scene go undetected even when the object undergoing the change is the focus of attention	Experimental Psychology (Grimes, 1996)
Attention to Visual Detail	Detecting visual changes in the environment	Experimental Psychology (Levin & Simons, 1997)
Attention to Detail	Covering all aspects with painstaking accuracy; Has qualities of both completeness and correctness and involves noticing the "little things"	Quality Control (Cappel et al., 2005)
Attention to Detail	Heightened attention to visual detail	Quality Control (Rusconi et al., 2012)

The two literatures that provide interpretations and measurement strategies for detail focus that are more relevant for our purposes are cognitive/experimental psychology and quality control. In cognitive/experimental psychology, research on detail focus has primarily been on visual attention; that is, how people process multiple inputs at once, complete sorting tasks, or detect changes across time in visual representations of different scenes (Austen & Enns, 2000; Baylis, 1994; Baylis & Driver, 1992; Grimes, 1996; Levin & Simons, 1997; Rensink, & O'Regan, 1997). The quality control literature – spanning across fields such as manufacturing (e.g., Alefari et al., 2018), information technology (IT; e.g., Cappel et al., 2005), and security/screening professionals (e.g., Merks et al., 2018; Rusconi, 2012) – is similar to the cognitive/experimental psychology literature in that it often examines visual cognitive abilities. Innate visual abilities, such as accurately distinguishing relevant objects from others in the background on an X-Ray machine, are assessed to differentiate high performers from low performers in these fields (e.g., Hättenschwiler et al., 2018; Rusconi et al., 2012). These types of

tasks are more like the task requirements for Army Soldiers and therefore a better representation of the type of measure that would be appropriate for this project.

Based on the information obtained regarding detail oriented, the decision was made to focus on visual stimuli for the development of this measure, as opposed to other sensory stimuli such as auditory. This decision was made as visual stimuli are most commonly studied in the literature and would be easier to present in a testing situation. Other factors that were considered in conceptualizing the measure were to (a) ensure that working memory was not required in order to complete the assessment, such as having to remember and then report details in a picture, and (b) contextualize the measure for Army situations, if possible, to provide face validity.

The two most relevant measurement strategies found in the literature were from the quality control and experimental psychology literatures. In the quality control literature, visual screening measures are used for selection and training for security screening, such as baggage screening at airports (Merks et al., 2018). In these assessments, realistic X-ray images are shown to test takers, who must detect specific items among the images (Merks et al., 2018). In the experimental psychology literature, test takers may be asked to complete a changing picture task, in which they must view two rapidly alternating versions of a photographed scene that differ from each other and detect the changing element(s) (e.g., Rensink et al., 1996). Another example from this field is a task that presents test takers with pictures comprised of multiple shapes and asks them to indicate a) the relative location of objects within a scene, and b) the relative location of parts within an object (e.g., Baylis, 1994). This focus on relative locations in the assessment, however, could provide a confound with spatial ability. The project team developed initial concepts for a side-by-side picture task, in which participants identify differences between two pictures that are presented next to each other, and a replication task in which participants would be given a detailed graphic or text document that they would need to reproduce from memory.

Measurement Strategy Recommendations

The SME review panel agreed with the proposed measurement approaches and recommended that (a) the assessment use pictures or maps as stimuli if possible, rather than written documents to minimize the verbal comprehension requirements of the assessment, and (b) potential overlap with other constructs be reviewed to determine if there was sufficient overlap to enable a combined measurement approach. The requirements described for Detail Focused and Precise activities were similar in nature to those proposed for the Processes Information and Data construct, which involved detailed coding, checking, and verifying. Upon review, the project team determined that, while there was a conceptual distinction between the two constructs, operationally they were inextricably linked. From a theoretical perspective, detail focus is a necessary requirement for processing information accurately. From an operational perspective, however, the measurement strategies for the constructs would be quite similar, suggesting that separate measures would be unlikely to show differences between the constructs. In addition, combining the two construct measures would result in both time and resource efficiency.

The decision was therefore made to combine the two constructs into a single cognitive construct called Detailed Information Processing, which is defined as “Attentive to detail and

thorough, accurate, and precise in processing (compiling, categorizing, auditing, and verifying) information." As described in the section on Processes Information and Data, the recommended measure asks the test taker to review a map or graphic or complete a form and respond to questions in order to assess their Detailed Information Processing capabilities (see Appendix B).

Multitasking

For the purposes of this project, Multitasking was defined in the ATAF as "Rapidly processes and prioritizes multiple demands simultaneously and takes appropriate action. Takes appropriate action when multiple things compete for his or her attention" (Royston et al., in preparation). Across the various Army attribute lists, there were multiple constructs related to Multitasking which can be seen in Table 5. The USMA list of 21 Talents includes a construct called "Multi-Tasker," which has the same definition as Multitasking from the ATAF. In the ARI list of 22 cognitive constructs, there are two constructs related to Multitasking: "Focus" and "Attentiveness." Focus highlights the ability to process multiple sources of information at the same time, and attentiveness emphasizes an ability to shift attention between activities.

Table 5

Army Constructs Related to Multitasking

Construct	Definition	Source
Multi-Tasker	Rapidly processes and prioritizes multiple demands simultaneously; Takes appropriate action when multiple things compete for his or her attention	USMA 21 Talents (U.S. Army Research Institute, 2019)
Multi-Tasking	Rapidly processes and prioritizes multiple demands simultaneously; Takes appropriate action when multiple things compete for his or her attention	Army Talent Attribute Framework (Royston et al., in preparation)
Focus	Mentally processes multiple sources of sensory information/data at the same time while avoiding distractions (e.g., flying a helicopter or commanding an armored vehicle)	ARI 22 Cognitive Constructs (U.S. Army Research Institute, 2019)
Attentiveness	Focuses on the problem or situation and shifts attention between activities when appropriate	ARI 22 Cognitive Constructs (U.S. Army Research Institute, 2019)

A review of the literature found examples of Multitasking constructs described across a variety of academic fields, as well as spanning a number of subdisciplines of psychology (e.g., neuropsychology, experimental psychology, computer and human behavior, etc.). A sample of definitions from the literature can be seen in Table 6. While there were many similarities in the definitions found across the literatures, there were also some slight differences in emphasis. For instance, neuropsychology places an emphasis on the allocation of attentional resources when faced with competing demands (e.g., Burgess et al., 2000) whereas fields such as computers in

human behavior (e.g., Carrier et al., 2009) and communication (e.g., Wang, & Tchernev, 2012) emphasize the act of combining one specific task (e.g., media use) with other tasks. One related construct that differs more significantly from the Army's conceptualization of Multitasking can be found in the experimental psychology literature. Known as "supertasking," this construct only occurs in a very small portion of the population and is defined as showing no performance decrements when engaging in dual versus single task conditions (Watson & Strayer, 2010). While some cognitive research interprets "simultaneous" to mean completing two tasks *at exactly the same time,* the project team decided that for the purposes of this project, given officer job requirements it would be better to interpret simultaneous to mean, *during the same timeframe*. Across these literatures, however, the definition of Multitasking is widely agreed upon and can be broadly summarized as "the ability to make conscious shifts of attention between two or more tasks" (Wetherell & Carter, 2013).

Table 6

Multitasking Definitions from the Literature

Construct Name	Definition	Source Literature
Multitasking	How attentional resources are allocated in the face of competing demands	Neuropsychology (Burgess et al., 2000)
Multitasking	Simultaneously combining media use with other tasks	Computers in Human Behavior (Carrier et al., 2009)
Multitasking	Complete multiple tasks concurrently; Attend and respond to several different stimuli simultaneously with varying levels of workload	Psychology (Wetherell & Carter, 2013)
Supertasking	Showing absolutely no performance decrements with respect to performing single and dual tasks	Experimental Psychology (Watson & Strayer, 2010)
Media Multitasking	Engaging in two or more media activities at once.	Communication (Wang & Tchernev, 2012)

Given the information from the literature, the project team began to conceptualize recommendations for measurement approaches. Two general measurement approaches for Multitasking have been identified (Hambrick et al., 2010). The first approach focuses broadly on multitasking situations, which are defined as situations in which the participant must consciously shift their attention between two or more tasks (Hambrick et al., 2010). These measurement approaches examine how well participants perform when presented with multitasking situations in which they are asked to complete multiple tasks at the same time (e.g., Colom et al., 2010; Delbridge, 2000; Konig et al., 2005). The second approach focuses on Multitasking as a time-sharing ability (Hambrich et al., 2010). Test takers are given measures to assess this time-sharing ability (e.g., Salvucci & Taatgen, 2008; Nijboer et al., 2013), sometimes paired with a battery of cognitive ability tests. In the time-sharing ability assessment, one task is designated a primary

task and the other a secondary task; for example, test takers need to read an essay for comprehension while writing down dictated words (Salvucci & Taatgen, 2008). Time-sharing is operationally defined as performance differences in single-task versus dual-task conditions.

Measurement Strategy Recommendations

Initial concepts were developed for a series of task-based assessments of Multitasking. In one concept, the test taker is presented with two or more tasks and given priorities regarding task completion. The test taker engages in the assigned tasks and must react as changes occur to the stimuli. The context could be contextualized for the Army or be random contexts. This concept would align with the multitasking situation approach, requiring test takers to shift their attention between multiple tasks. In another concept, test takers would be given a group of tasks to complete within a fixed period of time. Test takers would be given different points for completing different tasks and would be asked to maximize their overall score. This concept would align with the multitasking as a time-sharing ability approach, asking test takers to complete multiple tasks at the same time.

When the concepts were presented to the SME review panel, they agreed that a task-based approach similar to those in the literature would be useful, but also agreed with the project team that capturing both the prioritization aspect and the skill at accomplishing multiple tasks at once may be difficult within a single approach. It was determined that modifications should be made to the construct definition, given that the two different aspects – completing multiple tasks at once and prioritization – might be conceptually distinct enough that they would be better assessed with different strategies.

Given that the prioritization aspect of the definition would involve the use of judgement and decision making, and there were other constructs in the ATAF that involve judgement and decision making, it was decided that it would be best to focus the assessment of Multitasking on the attention shifting aspect of the construct. Based on these discussions, the working definition of Multitasking was therefore revised to, “The ability to shift attention between two or more tasks with minimal reduction in performance.”

The time-sharing ability approach was selected for the final measurement recommendation as this approach is highly relevant to the job situations in which officers work. Officers are frequently in situations in which they are engaging in a primary task (e.g., planning), while at the same time they need to field interruption tasks (e.g., questions from their staff). This type of situation mirrors the situations found in time sharing measures in the literature which designate a primary task and compare performance between single and dual task conditions (e.g., Brookings, 1990). A dual task measure was proposed for Multitasking that consists of a primary visual search task and a secondary interruption task. This measure is unique from those found in the literature in that it is contextualized for Army officers. In the primary task, test takers are asked to select Soldiers with certain characteristics for a special unit. They are provided with a list detailing Soldier characteristics and rules regarding Soldier selection. Test takers are given 10 seconds per item to find the appropriate Soldiers before the test advances to the next item. An initial set of 8-12 items will be completed without interruptions and will serve as the baseline for performance. After establishing a baseline score, test takers will complete another set of approximately 28 items, but this time will also receive interruption prompts every 3-5 items.

These prompts will be in the form of pop-up windows made to look like messages from a colleague or superior asking for information about a Solider (e.g., whether they are deployable) or general information about the set of Soldiers (e.g., whether there at least five deployable Infantry Soldiers in the current list). For these tasks it will be necessary to ensure the test taker is able to look at the list of Soldiers so that they do not have to rely on their memory to answer the questions. Scores will be based on the difference in search performance between the single and dual task conditions, using either a formula scoring approach that adds points for each correct click (hits) and subtracts points for each incorrect click (false alarms) or a sensitivity metric like d-prime which accounts for hits, misses, false alarms, and correct rejections. More detailed information on this measure can be seen in Appendix C.

Problem Solver

Problem solving is a ubiquitous cognitive process. A 'problem' occurs when there is a goal that is not immediately attainable, whether the goal is specific (e.g., figuring out why a car is making an odd noise) or general (e.g., choosing a career); simple (e.g., choosing what to make for dinner) or complicated (e.g., designing a behavioral intervention); well-defined (e.g., solving algebra problems) or ill-defined (e.g., writing a short story). Problems occur everywhere, all the time, and are exceedingly diverse. Problem solving as a cognitive process reflects that diversity in that it has many forms and varied components. Within the ATAF there are four problem-solving constructs: Problem Solver, Structured Problem Solving, Troubleshooting, and Unstructured Problem Solving. Definitions for each of these can be seen in Table 7. The focus for this effort was on developing an assessment for the Problem Solver construct; that is, having the *"*Capacity to choose between best practices and unorthodox approaches to reach a solution. Accomplishes the task." (Royston et al., in preparation).

A challenge in identifying relevant literature and assessment ideas was that the two sentences of the Problem Solver definition offer somewhat divergent directions for the literature review. The first sentence of the definition references a judgment component of choosing between different approaches. This implies that (a) a problem solver identifies different approaches for solving problems, (b) has the skills to execute those approaches, and in particular, (c) knows of, and is skilled in the best practice approaches for a particular domain. In addition, the focus on choosing between best practices and unorthodox approaches suggests relationships to other constructs such as critical thinking, reasoning, and judgment. On the other hand, the second sentence of the definition emphasizes an execution component that goes beyond knowledge, skill, and judgement. ″Accomplishes the task″ implies that the Problem Solver is unhindered by different problem types in reaching task goals. In addition to knowledge, skills, and judgment, the second sentence also reflects a motivational component in the form of persistence to overcome problems.

Table 7

Army Talent Attribute Framework Constructs Related to Problem Solving

Construct	Definition
Problem Solver	Capacity to choose between best practices and unorthodox approaches to reach a solution; Accomplishes the task

Construct	Definition
Structured Problem Solving	Analyzes readily obtained information and evaluates results to select the best solution from a set of existing approaches to solve a problem
Troubleshooting	Determines causes of operating errors and decides what to do about them
Unstructured Problem Solving	Identifies complex problems, gathers related information, evaluates information relevance, evaluates the credibility of alternative information sources, and generates alternative solutions

Note. These constructs and definitions are taken from the Talent Management Task Force (Royston et al., in preparation)

Although the construct under consideration is specifically named Problem Solver, the research literature most often used the term "problem solving," so in this discussion of research the term "problem solving" will typically be used. The initial review of the literature indicated that problem solving is a complex, multi-staged construct with numerous opportunities for measurement. Unfortunately, different measurement approaches have been used to measure the construct in different situations (e.g., see Davis, 1973) and there is lack of validity generalization findings across contexts (Bell, 1979; Scriven, 1980). While there are a variety of problem-solving tasks available in the literature, they are not generalizable beyond the laboratory and are seldom consistent from study to study – a condition that has earned problem solving the title of "most chaotic of all identifiable categories of human learning" (Davis, 1966, p.36). In general, consistent research methods and measurement approaches are not yet being used to study problem solving, with the exception of some specific types of problem solving, such as creative and complex problem solving (e.g., Funke, 2010; Neubert, 2016; Vernon, Hocking, & Tyler, 2016). Some empirical and theoretical progress is also being made by breaking down problem solving into component constructs which can be measured as antecedent constructs that predict problem solving (e.g., Chen, Hertzog, & Park, 2017; Mohaghegh & Furlan, 2020; Revlin, 2013).

The development of measurement tools that examine these antecedents to problem solving performance has allowed for further individual differentiation in problem-solving performance. Constructs such as critical thinking (Butler et al., 2012), inductive reasoning (Klix, 2001), metacognition (Schraw et al., 1995), and epistemic monitoring (Greeno, 1989; Kitchener, 1983) have all been proposed to predict problem-solving success. Measures of these constructs could potentially serve as proxies for predicting how well an individual would do in a problem-solving situation. However, no single antecedent can be used to perfectly measure problem solving. Because of this issue, multiple antecedents would be required in order to accurately approximate problem-solving ability if an antecedent measurement approach were used. Thus, the project team decided not to measure antecedents of problem solving, but rather to measure problem solving directly. In addition, the decision was made to not use self-report of motivation or dispositional characteristics and to measure skills (e.g., demonstrated ability to solve problems) rather than behavioral tendencies (e.g., biographical data). In measuring problem solving directly, however, the project team would need to avoid focusing on specific kinds of problems or specific kinds of problem solving (e.g., those that require particular contextual knowledge) in order to ensure that officers would not have an advantage due to differences in their prior knowledge or experience. This is a challenge because domain-specific knowledge and

skills are a central feature of problem-solving measures. Since most problems are context-specific, the pool of available measurement approaches is limited for general problem-solving knowledge or skills.

The literature notes many general problem-solving methods that are selected by individuals based on the problem itself or the problem environment. These methods may include the use of an algorithm or step-by-step procedure for moving from initial state to desired state, planning out a solution and executing the plan, or generating and testing feasible solutions in a trial-and-error fashion. Other problem-solving strategies include searching for a solution through a problem space, searching for partial solutions that merely reduce the problem (Newell & Simon, 1972; Nilsson, 1971), searching for the simplest solution that satisfies minimum requirements (satisficing), or the using heuristics to achieve fast and likely, though not guaranteed effective, solutions.

Research on generalized problem-solving capabilities suggests that there are problem-solving approaches that are context-agnostic but applying these approaches to a given problem still requires domain-specific knowledge (Larkin, 1980). For example, the general problem-solving approach to scientific inquiry is to state competing hypotheses, then collect data to find support (or fail to find support) for your predicted hypothesis. However, to apply this approach to a specific question, the inquirer would need to be able to identify meaningful hypotheses, understand what counts as evidence, and be able to interpret that evidence in light of the experimental context and the body of experimentation that came before. These steps require domain-specific knowledge, and that knowledge gives an expert in the field much better problem-solving capabilities than a novice. The role of knowledge in problem-solving seems to be a rate-limiting factor, which undermines the development of general problem-solving capabilities across and within fields. Teaching general heuristics does not seem to enhance general problem solving in fields such as mathematics (Schoenfeld, 1979) or physics (Larkin, 1980).

Furthermore, specific domains typically have their own recurring problems, and problem-solving processes and procedures are often domain specific. The development of expertise in a field comes, in part, from developing problem-solving strategies and tactics for the recurring problems within that field (Larkin, 1980; Chi, Glaser, & Rees, 1982). Expertise also comes from developing domain specific knowledge and schemata (Chase & Simon, 1973; de Groot, 1966), which are essential for solving specialized problems. Thus, specific knowledge plays a critical role in most real-world, on-the-job problem solving where specialists are tasked with solving recurring problems or problems of a distinct type.

Despite these issues, a general approach might still be the most appropriate way of solving problems in many cases. For example, Baird (1983) points out that some heuristics may be the most successful problem-solving approach when new types of problems emerge. Approaches such as working backward from a goal, means-ends analysis, and problem decomposition reoccur across contexts reflecting a degree of domain generality. However, Larkin (1980) notes that these general strategies are inefficient and tend to be applied when the problem solver has no other logical starting point. As an example, Simon and Simon (1978) compared the use of means-ends analysis in novices and experts in physics; they found that those without in-depth physics knowledge used means-ends analysis while subjects with in-depth

physics knowledge took a more direct approach to solving the problems and applied known physical representations and principles. Larkin (1980) argues that domain expertise comes, in part, from developing specific strategies that are most efficient for solving problems in that domain. In general, Newell (1980) called working backwards and means-ends analysis 'weak methods' because they start from a domain-naïve position.

One general approach that is often used by experts when dealing with ill-defined or novel problems is problem decomposition. With this approach, a problem is broken down into sub-problems or made simpler by abstraction. Breaking a problem down to an abstraction can be seen in many forms of diagramming and modeling where aspects of the problem can be stored on paper rather than in short-term memory. Architectural blueprints are a good example of modeling a hard problem with a simplified version. Solving architectural problems on paper is a much better strategy than addressing the problems while building (Larkin, 1980).

Beyond general problem-solving approaches, there are general problem-solving principles that may or may not be appropriate for any given situation but are important to be aware of. Rubenstein (1975) reviewed several principles used to avoid problem solving pitfalls. Larkin (1980, p. 24-25) summarized Rubenstein's principles as follows:

a. Avoid getting caught in detail--attempt to see the general pattern or picture; Go over the problem several times until a pattern develops
b. Withhold premature commitment to a single strategy--consider several before proceeding
c. Create models--verbalize, make graphs, write down, create abstract or concrete models. A model simplifies the problem
d. Find new representations of the problem; transform it into another system or context.
e. Question your premises and even reject them, and, if necessary, replace them with others or innovations
f. Verbalize your situation, ask questions, use different words. This may lead you to recall relevant information from your long-term memory
g. When the goal is specified, work backward to the beginning
h. Locate stable substructures that can serve as touchstones in the solution that you can branch out from and return to
i. Think of and apply analogies and metaphors; This places the problem into a larger structure allowing for solution
j. Incubate--leave the problem alone for a while and do something else; Talk to a colleague who may have a different viewpoint that could suggest a solution

Though there is some overlap in strategy among Rubenstein's techniques, they are labeled problem-solving principles and not strategies or approaches because they are not specifically tied to situations or uses (Cyert, 1980). As such, they may provide a useful foundation for obtaining a general assessment of problem solving. In addition, despite the variety and diversity that exists across various problems, some core characteristics of problems have been identified: goals, resources, constraints, and actions (Glass & Holyoak, 1986). First, the unattained *goal* provides a description of a preferred state. Second, a problem is contextualized in a situation with relevant *resources* or other objects that could be used to reach the goal. Third, the situation also provides *constraints* that limit the acceptable solutions or feasible approaches

to reaching the goal. Finally, given the resources and constraints, there are operations or *actions* that can be used to create the solution.

Problem solving is often viewed as a cyclical process that unfolds in a series of stages. Wirth and Klieme (2003) defined problem solving to include search and application of strategies to transform the current state to a desired state. According to these authors, the problem-solving process is comprised of five cognitive and metacognitive components: searching for information, planning the solution, taking action, processing feedback, and optimizing actions. This process is similar to other problem-solving cycles (e.g., Bransford & Stein, 1993; Hayes, 1989; Pretz et al. 2003; Sternberg, 1986). Sternberg (2009) provides a seven-stage model of problem solving: a) problem identification, b) definition of problem, c) constructing a strategy for problem solving, d) organizing information about a problem, e) allocation of resources, f) monitoring problem solving, g) evaluating problem solving. Many of these stages look forward or backward to other stages and some stages are skipped altogether. Similar models for decision making such as the Observe, Orient, Decide, Act (OODA) loop represent repeated decision-making cycles that contribute to solving a problem. The Military Decision-Making Process (MDMP), while not a cyclical model, is a recurring model where problems are identified, and solutions are generated and tested.

Perhaps the simplest model of problem solving is the 4-stage model described by Glass and Holyoak (1986). This model consists of initial problem representation, solution generation, solution execution, and feedback/evaluation stages. Most other models are variations on this model with more or less emphasis placed on unpacking one or more of these stages. However, it is important to note that these stages are not unitary cognitive processes in themselves but can be unpacked and expanded to show underlying information processing tasks involved. These cyclical stage models help to elucidate the variety of supporting behaviors and capabilities that enable a person to proceed from problem detection to solution. Each stage in the process can be viewed as a problem-solving competency, with enabling knowledge, skills, and abilities. While problem-solving cycle models help to elucidate the variety of supporting behaviors and capabilities that lead to solutions, they also highlight the role of the domain or context. The search for information stage of Wirth and Klieme's (2003) model points to the need to understand the situation in which the problem arises.

Measurement Strategy Recommendations

Several possible approaches to assessing the Problem Solver construct were presented to the SME review panel, such as SJT assessments that would present officers with a series of problems and ask them to select the best and worst tactics for solving the problems or the most important considerations or factors in the problems. The SME review panel indicated a preference to have the assessment focus on the core of general problem solving, avoiding specific types of problem solving such as creative or complex problem solving, and emphasizing structured problems as opposed to unstructured problems. In general, structured problems have one solution and if the Problem Solver knows the solution for a particular problem, they should be able to solve the problem. Therefore, the modified recommendations for measuring Problem Solver were to focus on generalized, structured problem solving with a knowledge-agnostic, skill-based task. This necessitated a change in the operational definition for Problem Solver. The

revised definition was, “Understands multiple approaches for representing and solving problems.”

Decomposing this definition into its constituent parts helps to emphasize the intended meaning. The term “understands” acknowledges that this trait is composed of a combination of knowledge and skills, and the phrase “multiple approaches” emphasizes breadth. An expert in one domain may proficiently solve problems for that domain, but not score highly on the general problem-solving attribute because they lack sufficient breadth in problem solving knowledge and skills to perform well outside of their domain of expertise. Further, the phrase “representing and solving problems” acknowledges that the construct is grounded in a basic information processing view of problem solving as a multi-staged process. Problem solving can be concisely summarized in four stages: a) form initial representation, b) use methods to form a plan, c) reformulate the plan as needed, d) execute the plan and test the outcome (Glass & Holyoak, 1986). The two initial stages are the most cognitively oriented, least tied to domain-specific procedures, and most feasible to measure with a question-and-answer format. These two stages are reflected as “representing” and “solving” problems in the definition, which are two proposed dimensions of the Problem Solver construct (Representations and Approaches). Definitions for these dimensions are:

- Problem Solver - Representations: Knowledge of and skill in forming external representations of an ill-defined problem
- Problem Solver - Approaches: Knowledge of and skill in identifying methods and plans for solving problems

The project team used these definitions to create the concept for a multiple-choice test of general problem-solving knowledge and skills that would consist of items measuring both dimensions. Scenario-based stems would be used to frame the questions before asking dimension-specific questions. For the *Representation* items, respondents would be asked to identify the best external representation of a problem given a goal (e.g., reframing, communicating, understanding the problem). Correct answers are verifiably correct, based on prescriptions in the literature for creating external representations and visualizations. For the *Approaches* items, respondents are asked to identify the best approach for solving a problem or making progress toward solving a problem. *Approaches* items may also deal with formulating a plan based on the chosen methods for solving the problem. Correct answers would be determined based on both prescriptions in the literature and SME consensus. Example items for each dimension can be seen in Appendix D.

Systems Thinking

Systems Thinking involves understanding patterns of interrelationships and is defined as “conceptualizes and understands relationships and arrangements within and between relevant components and structures” (Royston et al., in preparation). According to Senge et al. (1994) complex systems are comprised of elements which continually affect each other over time and operate toward a common purpose. Systems thinking skills enable leaders to recognize existing systems, identify patterns of interrelationships among key components of the system, and determine the impact of actions on the system.

Systems Thinking requires in-depth knowledge of the given problem domain, which may be obtained through the systems thinking process but often requires preexisting knowledge. As a result, Systems Thinking assessments often use ad hoc methods that are developed for narrow applications. For example, Davis and Stroink (2016) assessed Systems Thinking specific to ecology and environmental effects, Griffiths and Grant (1985) looked at Systems Thinking in an educational context, and Palaima and Skaržauskienė (2010) assessed Systems Thinking specific to manufacturing and retail trade industries. This domain-specific, knowledge-driven nature poses a problem for measuring Systems Thinking across individuals, as a test developed to measure Systems Thinking within one domain may not be applicable to another domain. In contrast, the assessment that is being developed for this project must be appropriate for all Army leaders regardless of their previous experience or assigned branch in the Army and should not require that the applicants have existing knowledge about a given type of system.

There are some Systems Thinking assessments that do not require domain-specific knowledge; often these assessments will use self-ratings or other ratings of Systems Thinking behavioral tendencies. For example, Moore et al. (2010) created 20-item self-report Systems Thinking scale (STS) that asks individuals to report systems thinking behaviors using a 5-point Likert scale. Example items are: "I recognize system problems are influenced by past events," and "I think small changes can produce important results." Factor analyses showed some indication of three factors in the STS: 20 items loaded highly on Factor 1 (System Interdependencies); 2 items also loaded on Factor 2 (Personal Effort), and 4 items also loaded on Factor 3 (Reliance on Authority). Randal and Stroink (2013) further developed the STS to create the Systems Thinking Scale Revised (STSR), which contained 15 items that used a 7-point Likert scale. Example items include: "Everything is constantly changing." and "If I make plans and control my behavior, I can accurately predict how my life will unfold (reverse scored)".

Skaržauskienė and colleagues (Kvedaravičius et al., 2009; Palaima & Skaržauskienė, 2010; Skaržauskienė, 2008; 2009; 2010; Skaržauskienė, & Jonušauskas, 2013) conceptualized systems thinking as a competency and measured it through self-report items from the education version of the Emotional and Social Competency Inventory[1] (ESCI- U; Boyatzis & Goleman, 2007). The items were divided into six dimensions based on evidence from an exploratory factor analysis: process orientation, systems logic, understanding of mental models, dynamic thinking, continuous learning, and interactivity. Each factor had a significant relationship with organizational performance, though most, except for process orientation, had small regression weights. In another analysis, this measure of System Thinking accounted for 32% of the variance in a self-report organizational performance measure (Skaržauskienė, 2010).

Palaima and Skaržauskienė (2010) used the ESCI-U to examine the importance of Systems Thinking factors for different types of leadership. As the authors expected, they found that more Systems Thinking factors were important for organizational/strategic leadership than for personal or relationship leadership performance. Dynamic thinking, interactivity, systems logic, and process orientation were important for organizational/strategic leadership, but only process orientation was important for personal/relationship leadership performance.

[1] Note that the University version of the ESCI includes two cognitive competency measures, systems thinking and pattern recognition, and is primarily used as a research tool for students and in the higher education setting.

Arnold and Wade (2017) developed a model of Systems Thinking that identifies relevant behaviors that could serve as a foundation for rating scales. The authors identified four overarching dimensions of Systems Thinking (Mindset, Content, Structure, and Behavior), as well as 16 specific skills within the four dimensions that provide a comprehensive model of systems thinking. The skills map onto different levels of systems thinking maturity and could potentially be used as the conceptual foundation for rating scales to measure Systems Thinking. Four skills from Arnold and Wade's (2017) framework are particularly relevant to the working definition of Systems Thinking: identifying relationships, characterizing relationships, identifying feedback loops, and characterizing feedback loops. At a basic level, these skills involve picking out and seeing relationships of varying levels of spatial or temporal separation and complexity and judging the strength and importance of those relationships.

Another potential approach for assessing Systems Thinking is to examine the words that writers use to describe the problems or systems. Booth-Sweeney and Sterman (2000) developed the Systems-Based Inquiry (S-BI) protocol, which analyzed the systems thinking level of students based on their responses to interview questions. The S-BI, designed to evaluate training effectiveness, includes five levels on a scale of systems intelligence, ranging from Level 0 (no response) to Level 4 (fuller utilization of systemic reasoning). It examines responses for indications of understanding concepts such as stocks and flows, feedback, and delays. Similarly, Grohs et al. (2018) developed a scenario-based instrument to measure Systems Thinking competency in a complex problem-solving context. After reading a community development scenario, test takers are given writing prompts, such as "Given what you know from the scenario, please write a statement describing your perception of the problems and/or issues…" These writing prompts are scored on dimensions of problem identification, information needs (information seeking), stakeholder awareness, goals and unintended consequences, implementation changes (foreseeing challenges), alignment (of stated problems, goals, and plans).

Using open-ended questions to assess Systems Thinking would be a challenge because they are often labor intensive to proctor and score and can introduce bias if not done under controlled conditions. However, using a systemic scoring rubric (e.g., Booth-Sweeney & Sterman, 2000, Grohs et al., 2018) and computer essay scoring techniques (e.g., Sherman & Burstein, 2013) would allow human scoring examples to train computer algorithms to categorize subsequent responses. In that way, an automated essay scoring tool could be developed to determine the level of an officer's knowledge based on the specific words and language that they use in their writing.

One final approach is to measure Systems Thinking skills through a simulation, such as the Beer Distribution Game (Watson, 2018). Originally developed by Forrester (1961) and refined by Sterman (1989), the Beer Distribution Game is modeled on a simple production and distribution system for a single brand of beer involving a retailer, a wholesaler, a distributor, and a manufacturer/factory. The player's goal is to maximize profit. There is a time lag between ordering and receiving the beer and the retailer and wholesaler do not communicate directly. The beer game demonstrates system characteristics such as the interconnectedness of elements, the complexity of relationships interacting to cause second or third order effects, and feedback loops. While the beer game was not developed and validated as an assessment tool, this kind of simulation would be suitable for measuring Systems Thinking skills.

Video games can also be used to provide the simulation content, like in Ecopolicy (Vester, 2010), an interactive game that was developed specifically to assess the player's systems thinking ability. The players' aim is to rule their country in an economical and ecologically sustainable manner under the premises of interconnected thinking. The players manage the state by investing in eight interconnected life areas: politics, production, environmental pollution, quality of life, sanitation, population, growth rate, and education. In the game, these areas are interlinked by mathematical relations that result in a chain of effects after a single decision.

One benefit of using simulations and games is that they can be used to provide fictional systems and information or opportunities for exploration to develop the required knowledge. This is the approach used by the Systems Thinking assessment currently under development at ARI (e.g., Adis et al., 2016). The ARI assessment takes the form of a game that elicits exploration behavior from the test taker, then tests how well that exploration leads to understanding and application of system knowledge. The player investigates test systems and then identifies the elements of each system, determines how elements are related, evaluates dynamic relationships within each system, and applies an integrated understanding to the overarching mission objectives. While the ARI Systems Thinking assessment game is potentially a very useful approach to assessing systems thinking, it currently takes multiple hours to complete and is in the process of undergoing reduction efforts, so it not practical for all testing situations as it stands currently.

Measurement Strategy Recommendations

Several possible approaches to assessing the Systems Thinking construct were presented to the SME review panel, such as video, graphic, or text-based scenarios that provide information about a system with multiple-choice or short answer assessment questions. The SME review panel emphasized the importance of focusing on methods that could be easily presented to test takers without complex hosting systems, easily scored without complex automated scoring, and assessed within a relatively short amount of time. Based on discussions with the SME review panel, the preferred approach for measuring systems thinking should involve customized multiple-choice items that describe the relationships between and among elements of a system.

In the final version of the recommended approach to assessing Systems Thinking, there are two proposed item types. In the first type, the test taker will read a text description of a fictitious system and select the best graphical representation of the relationships in the system. In the second type, the test taker is presented with a graphic model or image of a system and will be asked to select the multiple-choice response option that best describes the graphical model. The characteristics of the systems will be modeled after a variety of system types (e.g., biological, mechanical, social) of appropriate level of difficulty, but the assessment will use fictional content so that the test does not require any existing knowledge of specific systems. Example items are provided in Appendix E.

Perceptive

The ATAF defines Perceptive as "Effectively discerns the deeper meaning or significance of one's observations (e.g., events, people, and communication)"(Royston et al., in

preparation). One aspect of the definition worth noting is that it focuses on "one's observations." This suggests that a visual examination of some type of event or situation might be a useful stimulus for this assessment. Another aspect of note in this definition is the phrase "…discerns the deeper meaning or significance," which suggests that the individual will have some way of discerning deeper meaning from their observation; that is, some type of knowledge or contextual information that they can use to give meaning to their observation. The observation itself could be of any number of different events or situations. For example, if someone observes a group of people interacting and they have knowledge of nonverbal cues, that provides deeper meaning to their observation of the interaction. In the same way, if someone observes a house on fire and they have knowledge and experience with house fires, that will enable them to discern more meaning from their observation of the fire than someone who does not have knowledge about fires. The role of knowledge and experience in successfully discerning deeper meaning from observations poses a challenge for measuring the construct, given that one of our operating assumptions in the project is that the assessments should not require any specialized or military-specific knowledge or experience.

Multiple bodies of literature were examined as potentially relevant to this construct, including cognitive insight, social cognition, social and visual perception, and sense-making. Because social situations are one area for which there is a somewhat level playing field for knowledge and experience and given that some of the observation examples stated in the definition are specifically social in nature (i.e., people, communications), constructs such as social cognition and social perception were included in the review.

In a systematic review of cognitive insight, Van Camp et al. (2017) defined the construct as "the ability to re-evaluate thoughts and beliefs in order to make thoughtful conclusions" (p. 12). There are a variety of validated tools that assess this construct. One of the most prevalent is the Beck Cognitive Insight Scale (BCIS), a 15-item self-report questionnaire that assesses self-reflectiveness (measuring objectivity, reflectiveness, and openness to feedback) and self-certainty (decision making and resistance to feedback from others; Beck et al., 2004). Another tool is the insight orientation scale (IOS), which measures various aspects of insight including surprise (making connections), restructuring (change ability, conviction about a new meaning), level of consciousness (ability to experience feelings), problem solving (problem finding, problem shaping), complexity (abstraction, depth), self-reflectiveness (thoughtfulness), and awareness (attention, ability to perceive; Gori et al., 2015). These cognitive insight measures were not particularly useful in developing ideas to measure Perceptive given that conceptually cognitive insight is substantially distinct from our definition of Perceptive, as it focuses specifically on evaluating thoughts and beliefs. In addition, the cognitive insight measures are self-report rather than direct measures of skill or ability.

Another potentially useful body of literature that was considered was social cognition. Social cognition, or social inference/perception, is the ability to comprehend and interpret emotions, behaviors, and intentions of other individuals (McDonald et al., 2015). Given that all test takers would have some level of experience with social situations, no specialized knowledge would be needed for these assessments. The social cognition literature includes measures such as facial recognition, emotional expression/intelligence, theory of mind (ToM; the ability to recognize other people's mental states to understand and predict their behavior), and more complex social interactions.

There are a variety of tools that measure facial or emotional interpretation, such as the Bell Lysaker Emotion Recognition Task (BLERT) or Emotion Recognition Task (ERT; Pinkham, 2016), Emotion Bias Task (EBT; Pinkham, 2016), and Wechsler Advanced Clinical Solutions (ACS) Social Perception test (Kandalaft, 2021). The BLERT/ERT assess the ability to recognize emotion in facial expressions; an image of a face appears for a very short period of time and test takers must select which emotion the face was conveying (Pinkham, 2016). The Emotion Bias Task (EBT) is similar to the BLERT/ERT but limits the decision to two emotions. The ACS Social Perception subtest measures both facial affect recognition and affect recognition from prosody and interacting pairs of people (Kandalaft, 2021), with test takers completing tasks such as selecting an emotion to describe the person in the photograph or selecting a face that matches the emotion expressed by a speaker. Other tests focus on perceiving and understanding emotions, such as the Mayer-Salovey-Caruso Emotional Intelligence Test (MSCEI; Mayer et al., 2002), which uses scenarios typical of everyday life to measure how well people perceive, use, and understand emotions. While these measures do involve effectively discerning the deeper meaning or significance of one's observation, thus technically meeting the ATAF definition of perception, the project team determined that each of these was too focused specifically on facial and emotional recognition and did not meet the broader intent of the Perceptive construct, which included discerning the meaning of events and situations.

Therefore, the project team examined assessments of more complex social interactions. One tool that evaluates a broader range of social perception is the Awareness of Social Inference Test (TASIT; McDonald et al., 2015), which measures emotional recognition and social inferences. To measure social inferences, test takers are shown video clips depicting interactions and asked to respond to questions that interpret what they see. This type of measure specifically addresses the aspect of discerning deeper meaning regarding observations of people and communication. In another example, the Social Cognition Screening Questionnaire (SCSQ) measures verbal memory, schematic inference, ToM, metacognition, and hostility bias using verbal interviews (Roberts & Fiszdon, 2011). The test administrator reads short vignettes that present an interaction between a fictional character and the test taker, and then asks questions verbally about the vignettes to assess the test taker memory and understanding of the situation. This assessment also requires that the test taker discern deeper meaning regarding observations of people and communication. The concept of presenting these types of stimuli would be potentially useful for assessing Perceptive, although the individual nature of the test administration for the SCSQ is not a useful model to consider for assessment in the Army.

Another potentially relevant tool is the Edinburgh Social Cognition Test (ESCoT; Baksh et al., 2018), which measures four social cognitive abilities: cognitive ToM (i.e., the ability to recognize other people's mental states to understand and predict their behavior); affective ToM (the ability to recognize and predict other people's emotional states); interpersonal understanding of social norms and intrapersonal understanding of social norms. The ESCoT consists of a series of 11 social interactions depicted in frames of four cartoon-style picture animations: one practice item, five items involving a social norm violation, and five portraying everyday interactions that do not involve social norm violations. Each cartoon series has a different context and specific questions relating to that context. Baksh et al. (2018) presented the animations on a computer screen and at the end of each animation, a static storyboard depicted a summarized version of the interaction. Test takers were then asked questions to assess each of the four subtests of social cognition. This approach to having test takers "observe" an event or situation seems potentially

useful for measuring Perceptive and provides more complex situations than measures of facial and emotional recognition. In addition, it could be administered in a standardized format, either via a computer-based platform or a paper-based assessment tool. One advantage of the ESCoT is that overall performance was not related to measures of verbal comprehension and perceptual reasoning performance.

Measurement Strategy Recommendations

Several initial concepts for assessing Perceptive were proposed, such as the presentation of social scenes using video clips, cartoon frames, or text-based presentation of scenarios, followed by questions asking about the interpretation of social events or interactions in the scenes. While the approach to presenting the stimuli and responses was viewed positively, there was concern from the SME review panel that the measure was too focused on social cognition rather than a more cognitive task. The project team determined that a possible solution would be to maintain the format for presenting situation/event stimuli, but to target content that would measure whether people can discern a deeper meaning from non-social or less social aspects of what they are observing. The challenge was to devise scenarios where the test taker could discern a deeper meaning without (a) requiring any specialized knowledge that would give some test takers an advantage over others, or (b) creating a confound with test takers' ability to remember or capacity for attention to detail. Test takers would need to be directed to examine global content rather than specific details in a scene to prevent attention to detail from being a confound; that is, the test taker should be able to succeed not by attending to specific details but by extracting meaning from the picture.

The project team used research on assessments of "thin slices of behavior" to provide ideas regarding samples of behavior upon which test takers could make judgements (Ambady et al., 2000). In this research, test takers are shown brief videos lasting from a one to five minutes and must make ratings/judgments about the person being shown (Ambady & Rosenthal, 1992). A final measurement approach for Perceptive was developed that first provides test takers with written context regarding a situation, then displays a photo or graphic of a scene from which the test taker needs to discern meaning. Test takers respond to two types of items: (a) items related to tasks identified in the scenario, and (b) items related to social aspects of the scenario. Test takers use a 5-point Likert scale to rate the likelihood of various statements or respond to multiple-choice questions. Example items can be seen in Appendix F. Items for the scale could be related to military scenes or events but would not require prior military knowledge.

Interdisciplinary

The construct Interdisciplinary is defined as "synthesizes and applies knowledge from multiple disciplines into a coherent overarching perspective" (Royston et al., in preparation). An initial examination of the literature for this construct spans a number of disciplines (e.g., higher education, psychology, medicine, special education), and each focuses on slightly different aspects or interpretations of the construct. A sample of definitions from the literature can be seen in Table 8. In the higher education literature, where this construct is perhaps most present (e.g., Borg & Borg, 2001; Johannes & Kasteren, 1996; Klein, 1999; Klein, 1990; Mansilla et al., 2000; Newell & Green, 1982; Sill, 1996; Stowe et al., 2002; Vincenti, 2001), the focus is on effectively

integrating information from multiple disciplines in order to produce a cognitive advancement (e.g., solve a problem or create a product) that could not have been created otherwise.

Table 8

Interdisciplinary Construct Definitions from the Literature

Construct Name	Definition	Source Literature
Interdisciplinary Thinking	Emphasizing one or more academic traditions; Telling stories or analyzing data from multiple perspectives; Using integrated thinking	Higher Education (Lawrence, 2011)
Interdisciplinary Teamwork	Members of multiple disciplines working together to respond to social issues in fieldwork (e.g., repercussions from natural disasters)	Higher Education & Economics (Mendenhall & Berge, 2010)
Interdisciplinary Thinking	Inquiries which critically draw upon two or more disciplines, and which lead to an integration of disciplinary insights	Higher Education (Boix Mansilla et al., 2000)
Transdisciplinary Health Care	Transdisciplinary health care involves reaching into the spaces between the disciplines to create positive health outcomes through collaboration	Higher Education & Medicine (Nandiwada & Dang-Vu, 2010)
Interdisciplinary	The integration of disciplinary perspectives	Medicine (Newell & Green, 1982)
Biopsychosocial Model	Using biological, psychological, and social explanations to approach medical and psychological treatment of patients	Medicine (Engel, 1977)
Multidisciplinary Treatment Teams	The cohesive collaboration (effectively communicating and coordinating treatments for Autism Spectrum Disorder) across service professionals	Medical research (LaFrance et al., 2019)
Transdisciplinary Research	Cross-disciplinary collaboration among health scientists that bridges across multiple disciplines; Meant to foster collaborative research and training that ultimately leads to public health improvements targeting complex public health problems (e.g. cancer and obesity)	Community Psychology (Hall et al., 2008)

The psychological literature has a similar interpretation of the construct, defining Interdisciplinary as multiple disciplines working together to solve social issues (e.g., interracial conflict) in fieldwork (e.g., Hennessy, & Walker, 2011; Stokols, 2006). Similarly, the medical (e.g., Engel, 1997; Hall et al., 2008; Zaretsky, 2007) and special education (e.g., LaFrance et al., 2019; Ruddy & Rhee, 2005) literature both focus on integrating multiple disciplines and ways of thinking in order to address a problem. Within these literatures the terms “biopsychosocial” and “transdisciplinary” are sometimes used as synonyms for Interdisciplinary. The biopsychosocial

model is referred to in the medical literature and involves integrating multiple frameworks (biological, psychological, and social) to properly diagnose and treat patients (Nandiwada, & Dang-Vu, 2010). The transdisciplinary approach is mentioned in both the special education literature and the medical literature at large; it specifically refers to integrating multiple types of care providers in order to treat the student or patient in a more comprehensive manner (Mendenhall, & Berge, 2010). This approach can be aimed at a broad issue such as more effective primary care, or a more specific issue such as treating a complex disorder like autism.

The National Institute of Health (NIH) has a transdisciplinary medical research initiative that broadly aims to foster collaborative research and training by supporting organizations with improved collaborative-readiness factors (Hall et al., 2008). The NIH lists two types of collaborative-readiness factors that are important to foster transdisciplinary research; contextual-environmental conditions (e.g., institutional resources and supports or barriers to cross-departmental collaboration) and interpersonal characteristics (e.g., research orientation, leadership qualities) (Hall et al., 2008). Although the current research is focused on individual interdisciplinary characteristics, it may be useful in the future for the Army to consider that there are environmental conditions that affect the likelihood of effective interdisciplinary collaboration taking place.

Another approach to examining Interdisciplinary is to look for antecedents that allow individuals to work well together or understand diverse topics. With regard to the former, Klein (1990) has argued that interdisciplinary individuals need to be reliable, flexible, patient, resilient, sensitive to others, risk-taking, thick skinned, and have a preference for diversity. Ideally, the interdisciplinary person shows tolerance for ambiguity, takes initiative, shows a degree of assertiveness, and has a broad education (Klein, 1990). Other researchers have pointed to the need for cognitive capabilities such as critical thinking, creativity, problem solving capabilities, contextual understanding, and the cognitive and emotional capabilities to deal with complexity (Stowe et al., 2002). These capabilities allow the interdisciplinary person to make connections across disciplines, reason from analogy, or integrate cross-disciplinary ideas to reach a better understanding.

Related characteristics to the cognitive precursors of interdisciplinarity are integrated thinking (IIRC, 2013) and personal complexity (e.g., Schroder et al., 1967; Streufert & Nogami, 1989; Suedfeld et al., 1992). The International Integrated Reporting Council (IIRC) defines integrated thinking as "the active consideration by an organization of the relationships between its various operating and functional units and the capitals that the organization uses or affects" (IIRC 2013, p. 2). The purpose of integrated thinking, according to Dumay et al., (2018), is to reduce organizational siloes and improve internal efficiency. While integrated thinking emerged from, and is tailored to, business organizations, it has direct relevance to Army units and operations.

Personal complexity variables include conceptual complexity (Schroder et al., 1967), integrative complexity (Suedfeld et al., 1992), and metacomplexity (Streufert & Nogami, 1989). Generally speaking, these constructs reflect stable personality or cognitive style traits concerned with how individuals interact with complex environments. Conceptual complexity arises from recurrent differentiation of social and physical stimuli, and the subsequent integration and synthesis of ideas that lead to a nuanced perspective (Harvey et al., 1961; Schroder et al., 1967).

Integrative complexity, similarly, recognizes the role of differentiation and integration in fostering a complex mindset, but extends conceptual complexity to include social, physiological, or environmental factors (Suedfeld & Tetlock, 1977; Suedfeld et al., 1992). Metacomplexity theory is a more recent iteration of conceptual complexity that examines complexity at three levels of functioning: individual cognitive complexity, the interaction of cognitive complexity and other individual characteristics (process complexity), and higher-level impacts such as group or organizational complexity (Streufert & Nogami, 1989). These constructs, while not redundant with the individual interdisciplinarity construct, seem like essential precursors for success at engaging in interdisciplinary efforts.

There is a small body of literature that describes behaviors that interdisciplinary individuals or individuals working on interdisciplinary problems must enact (e.g., Newell, 2001; Klein, 1990). Newell (1998) rejects the notion that the interdisciplinarity construct changes from one profession/discipline/domain to another, but rather, suggests that an underlying process is involved in interdisciplinarity behavior across situations. Similarly, Klein (1990) outlined several universal steps to being interdisciplinary, including: defining the problem, determining knowledge requirements, developing an integrative framework, specifying what information still needs to be obtained, engaging in teamwork, searching for and gathering new knowledge, resolving conflicts, communicating, evaluating contributions, confirming and integrating information, and deciding directions.

Klein and Newell (1997, see also Newell, 2001) expanded the steps of the interdisciplinary process to include a phase of single-discipline analysis (composed of defining, determining, developing, gathering, studying, and generating) and inter-discipline integration (composed of identifying conflicts, evaluating assumptions, resolving conflicts, creating common ground, constructing shared understanding, producing a model, and testing that model). These broad behavior sets read like competencies and belie a multi-dimensional, multi-faceted process requiring many capabilities. Additionally, these models resemble models of other complex processes such as problem solving or systems thinking.

After examining existing constructs within the literature, a number of internal decisions and considerations were made about the direction with which to proceed with the construct. Firstly, it was determined that the next course of action was to decide what would be considered "disciplines" within the Army, as this would inform the development of the assessment. MOS, unit, and degree field (as this project deals directly with officers who must hold a bachelor's degree) were all considered as viable options within which the culture and knowledge pool might vary significantly to be considered "interdisciplinary." Another consideration was the application of the construct to officers throughout their careers. Upon reflection it seemed the Interdisciplinary construct would be more applicable to higher ranking Soldiers. In the literature, the construct is used to solve complex problems when a simple, uni-disciplinary approach will not result in an adequate solution. In the Army, junior officers are much less likely to be faced with complex problems, so the decision was made that a measure of this construct would need to target higher ranking officers rather than junior officers.

This construct was unique in that while many disciplines agree that an interdisciplinary approach can be very useful, there exists only one measure that could be found in the literature. This measure, part of a multidisciplinary class within an engineering school, presents

multidisciplinary groups of students with broadly formulated sustainable development problems as part of a class project. The project is then graded by the professor on a standard grading scale. Aside from this measure, it seems that employees and students are often encouraged to engage in a multidisciplinary approach to complex problem-solving and then thrown into an interdisciplinary team to sink or swim, as it were. Evidence of selecting for traits that could be associated with a propensity towards an interdisciplinary approach, perhaps another effective way to test for the same construct, was not found within the literature.

Measurement Strategy Recommendations

Three potential measurement approaches for Interdisciplinary were presented to the SME review panel. These included a SJT examining interdisciplinary behavior, an SJT examining interdisciplinary propensity, and a multidimensional forced choice test examining behavioral tendencies towards an interdisciplinary approach. In order to develop an assessment of interdisciplinary propensity or ability, a complex problem that required an interdisciplinary approach would be required. The project team discussed suitable types of complex problems within the Army with the SME review panel and several ideas were generated, such as mental health, medical, and complex war problems. The SJT examining interdisciplinary behavior was identified as the most desirable option of the three. Considering that Interdisciplinary is most useful to senior officers and could take a very long time to administer due to all the background information necessary to set up the complex problem, the decision was made to exclude this construct from further development.

Technologically Adept

The ATAF definition of Technologically Adept is “Learns how to use and apply advances in technologies or technological systems” (Royston et al., in preparation). The phrase "Learns how to..." in the definition suggests that acquisition knowledge and skills pertaining to technology are important components of this construct, and the phrase, "to use and apply" suggests practical knowledge and the ability to put that knowledge into action. The term "Adept" suggests that this construct also has a proficiency component. An initial examination of the literature indicated that, while there are many aspects of using and acquiring technologies that have been researched, there is no specific construct called “Technologically Adept” on which to focus a review of the literature. Instead, various precursors and related constructs were examined from extant literature and weighed as appropriate for the assessment of Army officers.

One area that has been well studied is the willingness to adopt and use technology. The Technology Acceptance Model (TAM; Davis, Bagozzi, & Warshaw, 1989; Venkatesh & Davis, 2000) indicates that two factors affect attitudes towards, and intentions to use, technology: the perceived ease of use and the perceived usefulness of the technology. Other related approaches to examining attitudes toward technology use include the Quality-Value-Satisfaction chain (Cronin, Brady, & Hult, 2000) or the technology-task fit paradigm (Goodhue, 1998; Goodhue & Thompson, 1995). While potentially relevant as an antecedent, Technologically Adept needs to focus on the cognitive ability or skill of learning how to use and apply technologies.

There are also many individual difference constructs that are potentially relevant to Technologically Adept, such as motivational (e.g., agency beliefs, self-efficacy), knowledge-

based (e.g., literacy), skills-based (e.g., proficiencies), and composite constructs (e.g., competencies). Some of the constructs that have been proposed are technological proficiency (Giacomimi, Lyle, & Wynn, 2012), technology self-efficacy (McDonald & Siegall, 1992), computer literacy (Merritt, Smith, & Renzo, 2005), computer self-efficacy (Compeau & Higgins, 1995), digital competence (Ferrari, 2012), web-oriented digital literacy (Hargittai, 2005), internet self-efficacy (Eastin & LaRose, 2000; Joyce & Kirakowski, 2014), and internet skills (van Deursen, Helsper, Eynon, 2014, 2016; van Deursen & Van Dijk, 2009a, 2009b, 2010, 2011; van Deursen, Van Dijk, & Peters, 2012; Van Dijk & van Deursen, 2014).

A key component in whether someone adopts and uses a technology or tool is the belief in their ability to use that technology effectively; that is, their efficacy beliefs about their technological ability. Someone who does not believe they can use the technology effectively (i.e., has low efficacy beliefs) is unlikely to use that technology. Because personal efficacy beliefs are a central motivating factor in the decision to use technology, they are important precursors in the development of technological knowledge, skills, and competencies. Self-efficacy is the belief in one's capabilities to perform specific tasks (Bandura, 1977). Applied to the domain of technology use, this could entail generalized beliefs about technology or 'technology self-efficacy'; the belief that one can successfully perform a new task using sophisticated technology (e.g., McDonald & Siegall, 1992). McDonald and Siegall examined this construct using five items and a Likert-type scale and found that technology self-efficacy is broadly related to numerous job outcomes, including job satisfaction, organizational commitment, work quality, and attendance. Efficacy beliefs can also be examined in the context of more specific technologies such as computers or the internet (e.g., Compeau & Higgins, 1995; Eastin & LaRose, 2000). More recently internet self-efficacy measures have been developed to address changing technology and population proficiency levels (e.g., Joyce & Kirakowski, 2014).

Related to the construct of self-efficacy are constructs such as technophobia (Frideres et al., 1983; Heinssen, Glass, & Knight, 1987; Rosen, Sears, & Weil, 1987; Rosen & Weil, 1990a, 1990b), computerphobia (Jay, 1981; Rosen, Sears, & Weil, 1993) computer anxiety (Raub, 1982), technostress (Brod, 1984), and computer aversion (Meier, 1985; 1988). These constructs reflect the lack of self-efficacy for technology or self-efficacy for computers. Unfortunately, there is little consensus on the boundaries between the various constructs and terms have been used inconsistently (Korukonda, 2005). While the self-efficacy and lack-of-self-efficacy measures are important for learning and applying technology on the job, most measures of efficacy beliefs are self-report measures, so are not ideal for measuring Technologically Adept, given the assumption that we are striving to use direct measures of cognitive skill or ability rather than motivational and/or self-report measures.

Another relevant approach to measuring aspects of the Technologically Adept construct would be to focus on skills (van Deursen et al., 2014, 2016; van Deursen & Van Dijk, 2009a, 2009b, 2010, 2011; van Deursen et al., 2012; Van Dikj & van Deursen, 2014), proficiency (e.g., Giacomini et al., 2012), competency (e.g., Ferrari, 2012), or literacy (Hargittai, 2005; Merritt et al., 2005). The literature tends to focus on either computer technology or internet technology when researching these types of constructs. There is often little that clearly distinguishes between skills, proficiencies, competencies, and literacy, and no specific guidelines are provided for what specific technologies would be included if Technologically Adept was assessed using tests of

these sub-dimensions. Therefore, additional research and consideration would be needed regarding which technologies would be appropriate content for this kind of measurement tool.

Measurement Strategy Recommendations

As stated at the outset of the report, one assumption for developing these assessment concepts was that they should focus on measuring the cognitive skill or ability directly, rather than using self-report or survey measures. Another stipulation identified specifically for this construct was to avoid developing an assessment that required knowledge of specific technologies or skills associated with specific technologies. This will ensure that officers with specific knowledge or experiences do not have an advantage in completing the assessment.

Given these two restrictions, the project team recommended using a performance-based task that leverages an invented technology to avoid creating bias. This fictional technology would be used to directly assess officers' ability to learn how to use and apply advances in technologies, as described in the construct definition. Test takers would be given an opportunity to familiarize themselves with an invented technology and demonstrate their performance using that technology. Additionally, upgrades to the technology could be incorporated so that the cycles of learning and becoming proficient repeat numerous times, requiring the test taker to demonstrate continuous adaptability to the changes in the technology. The advantage of this approach is that it can be used to examine actual use of technology, rather than the test takers' beliefs about their capabilities. With repeating update and learning cycles, this approach can also be used to measure learning and reflect that aspect of the definition. The disadvantage of this approach would be expected difficulty in creating a domain-free technology that goes beyond surface-level differences with existing technology. As such, this measurement approach would be more experimental and require multiple iterations of development and validation.

Other possible measurement approaches could include assessing antecedent measures that predict the likelihood of being technologically adept without directly measuring this construct. The literature suggests that certain antecedents such as technology efficacy beliefs (McDonald & Siegall, 1992), epistemic beliefs (Greene, Azevedo, & Torney-Purta, 2008), or exploratory behavior (Lyons, 1986) might successfully predict the Technologically Adept construct. The antecedent approach would measure the construct in a piecemeal fashion, however, with no single measure reflecting the entire construct. Another approach could be to use an essay or short answer format; the assumption being that people high in the Technologically Adept construct would write differently about technology than people low in the construct, such as using particular technology-related words or phrases. This assumption seems possible, yet no essay-based measures are currently presented in the literature.

After discussion of the literature and assessment options, the SME review panel recognized that it could be potentially beneficial to align the construct with the technophobia/ computer anxiety conceptualization. The Army would benefit by identifying officers who are not afraid of new technologies or computer updates (e.g., people who were not afraid of the transition to Microsoft 365). While this direction would be feasible, it is more of a dispositional trait, meaning it would typically be assessed with a self-report format and does not focus directly on a cognitive skill or ability. It would be possible to avoid using a self-report format if test takers were given a new invented technology for the purpose of the measure and assessed on

their exploration of and learning about the technology. Learning, performance, or even a physiological stress response could be measured while the test taker became familiar with the new technology. However, given the particular challenges faced with assessing this construct, the decision was made to exclude this construct from further consideration or development.

Discussion

Operational measures were investigated that relate to 10 cognitive constructs identified as important for officer success in the Army. In some cases, constructs were identified in the literature that were conceptually well matched to the defined Army construct, and in other cases, the constructs that appeared in the literature were only marginally related to the Army construct. One of our operating assumptions at the outset was that changes to the construct names and definitions should be avoided or minimized; in the end, however, recommendations were made to change the name and/or definition of six of the constructs for a variety of reasons. The key findings and recommendations for each construct will be discussed, as well as some general observations and recommendations for assessing the cognitive attributes of Army officers.

Analytical Thinking/Analyze Data or Information

When examining the literature, the terms, "Analytical Thinking" and "Analyze Data or Information" were not commonly found as construct names, so their associated definitions were used to identify constructs and measures that were potentially related. For both constructs, literature related to data analytics and Information Processing were commonly found. The definition of Analytical Thinking focuses on analyzing information and applying rules and logic, and the definition of Analyze Data or Information focuses on identifying principles by breaking down information. Both concepts are conceptually related to the domain of reasoning, which involves deriving or inferring conclusions based on existing evidence. There are three types of reasoning in the literature (inductive, deductive, and abductive), and while all types of reasoning were conceptually related to the analytical constructs, Army measures have already been developed for inductive and deductive reasoning, so our consideration focused specifically on abductive reasoning.

Abductive reasoning involves deriving or inferring conclusions as best or optimal based on incomplete or uncertain information. From an application standpoint, abductive reasoning is the type reasoning officers will use on a daily basis to make decisions, and solve problems, so having a tool to assess abductive reasoning would be highly relevant and useful. While the concepts for the two original constructs, Analytical Thinking and Analyze Data or Information each focus on unique aspects of analyzing information at a micro level, from a measurement standpoint capturing those distinctions may be difficult or impossible without a lab setting and, in addition, measuring the distinction may provide no practical benefit.

Given that a validated measure of abductive reasoning was identified that can serve as a model for development, there is a high likelihood of success in developing an operational measure of abductive reasoning for Army officers. The assessment approach developed was to provide brief scenarios with partial information and have test takers determine the most reasonable conclusions that can be drawn from the information. One aspect of analyzing information that was discussed as relevant to Army officers but was not captured in the validated

measure is an analysis of whether there is sufficient information with which to make a decision. This was identified by the SME review panel as a critical element of analysis for officers and therefore was integrated into the proposed measure.

Detailed Information Processing

Detailed Information Processing is the construct proposed to capture two of the original constructs, Processes Information and Data and Detail Focused and Precise. On the surface these two constructs may seem too conceptually distinct to combine, with the definition of Processes Information and Data focusing on skills across multiple processing tasks (compiles, codes, categorizes, etc.), and the definition of Detail Focused and Precise focusing on being thorough, accurate, and precise in completing a task. Often detail orientation is measured as a dispositional characteristic, but in the current project we are focused on constructs from the perspective of cognitive skills and abilities. Given this type of cognitive orientation, our focus with detail orientation was to assess the officers' skills at paying attention to details across various cognitive tasks. Given that Processes Information and Data is focused on examining officer skill across various cognitive tasks, combing the two measures was both reasonable and efficient. Success in tasks such as compiling or coding information will require both the capability to understand and complete the mechanics of the task and sufficient attention to detail to do it correctly and avoid errors. The assessment approach developed was to provide test takers with brief processing tasks that align with the set of actions in the Processes Information and Data definition (e.g., compiling, coding, categorizing, etc.) and to test participants' ability to perform those tasks accurately. The military-relevant tasks that were designed to measure the various Information Processing skills provide a highly face-valid test and could use maps or graphics as stimuli to reduce any overlap with verbal ability or reading comprehension.

Multitasking

Identifying a feasible assessment approach for Multitasking proved to be challenging due to vague and competing facets in its definition. Several aspects of the definition needed to be clarified or modified, including the meaning of "simultaneously," the role of "takes appropriate action," and the competing importance of *processing* multiple demands versus *prioritizing* multiple demands. Although some cognitive research is specifically interested in test taker performance of tasks that are literally being conducted at the same time, from a practical standpoint, officers tend to be in the position of needing to shift quickly from one task to another in the same period of time, not conducting the tasks at specifically the same moment. For this reason, "simultaneously" was removed from the construct definition.

The other aspects of the definition – prioritizing and taking appropriate action versus processing multiple demands – are both relevant to officer job requirements; however, they reflect different skills and abilities that may not be highly correlated and would be measured in different ways. Processing multiple demands requires being able to switch from one task to another without losing focus or making performance errors. On the other hand, prioritizing and taking appropriate action are decision making tasks that require critical thinking and analysis of options. While both aspects are important, the decision was made to focus on processing multiple demands rather than prioritizing, which is more commensurate with the definition of multitasking found in the literature. The assessment approach developed was to have test takers

complete a visual search task to establish a baseline score, then complete the same task while interruptive prompts are occurring to measure the level of performance decrement due to shifting tasks. Given the importance of prioritizing to managing and leading, however, it could be useful in the future to develop an assessment of prioritizing.

Problem Solver

The initial challenge with measuring Problem Solver was to identify the specific aspect of problem solving that would be useful and feasible to measure, given the parameters that problem solving skill should be measured directly, not use self-report or dispositional measures, and test takers should not be required to have any domain-specific knowledge to succeed. The original definition that was provided for Problem Solver was particularly complex because it focused on being able to choose between "best practices" and "unorthodox approaches" in order to reach a solution. This indicated that in order to succeed, the test taker would need knowledge in a given domain as well as an understanding of both well-defined and ill-defined creative problem-solving approaches that could be leveraged to solve a problem in the domain. After consultation with the SME review panel, it was determined that, for the purposes of this project it would be preferable to focus the assessment on generalized, structured problem solving with a knowledge-agnostic, skill-based task. This led to a significant change in the operational definition to focus the Problem Solver construct on understanding multiple approaches for representing and solving problems. The assessment approach that was developed uses a multiple-choice test format with scenario-based stems that frame questions about how to best represent and approach the problems described in the scenarios. This approach will be relatively simple to develop and administer and provides a useful complement to more complicated problem-solving measures that ARI has already developed to measure complex problem solving (see Adis et al., under review) and creative problem solving (see Adis et al., 2017).

Systems Thinking

Recommending an assessment approach for systems thinking was similar to problem solving, in that ARI already has an existing measure of systems thinking in development (e.g., Adis et al., 2017), so it was important to develop an assessment of systems thinking that provided a useful complement to the existing measure. Like problem solving, systems thinking is a highly complex construct that requires in-depth knowledge of a given problem domain as well as a set of multiple cognitive skills; in this case, skills that enable officers to recognize key components of systems and identify and understand patterns of relationships among the elements. The existing ARI systems thinking measure uses a first-person videogame context to provide test takers with an opportunity to learn about a series of fictitious systems on an alien spacecraft and respond to questions regarding the system elements, relationships, and application to mission. The existing measure provides a comprehensive assessment of multiple aspects of the systems thinking construct, but it also takes a few hours to complete and has certain computer hosting requirements. One objective with the current project was to develop concepts that were easy and quick to administer and score, though still measuring the actual skills or abilities directly rather than through self-reports of behavior. The proposed measure will capture one targeted piece of the broader construct of systems thinking, focusing specifically on understanding relationships between components and structures. The proposed items ask the test taker to select the static image or graphic depicting a fictitious system that is appropriate given a

text description of the system (or vice versa). While the measure provides a very targeted rather than comprehensive assessment of systems thinking, it will provide insight regarding officers' skill at understanding relationships in a system, using a quick and easy-to-administer format. This offers a useful and practical approach to measuring systems thinking as part of a broader assessment of officer cognitive capabilities.

Perceptive

Researching the construct Perceptive posed a challenge due to the existence of similar constructs or terms that are used across diverse bodies of literature. The term "perception" is used to represent an entire field of psychology that studies the interpretation of information from the senses. This body of literature is only peripherally related to the Perceptive construct, given that the Perceptive definition focuses on discerning the deeper meaning or significance of one's observations. Another term commonly found in the literature is "social perception" or "socially Perceptive." This body of literature is conceptually related to the Perceptive construct, as it seeks to understand how people discern the meaning or significance of social interactions. However, this approach effectively narrows the focus of the construct to social observations.

Determining a measurement approach for this construct therefore required focusing on specific elements of the construct definition, effectively discerns the deeper meaning or significance of one's observations. A central issue in the definition is understanding more specifically the actions that allow individuals to "discern deeper meaning" from something. Discerning deeper meaning from an observation is possible if the individual is able to make connections between what they observed and a broader understanding of elements of the situation or event. Having a base of knowledge or experience about a situation, event, or anything that would be observed would provide the individual with the potential for developing a deeper understanding of what they observed. This would be true regardless of whether they were observing social interactions, a sports game, a house fire, or clouds in the sky.

This poses a challenge for determining a measurement approach for the construct, however, because one of our operating assumptions is that no pre-existing knowledge should be required for the assessments. The measure we proposed for Perceptive is innovative in attempting to assess the construct outside the domain of social interactions, yet still providing a measure that does not require specific pre-existing knowledge. The assessment approach that was developed was to provide test takers with a text scenario, then display a scene from which test taker need to discern meaning. The meaning that must be derived reflects impacts to task objectives as described in the scenario or general, socially-oriented observations of the scene. From a measurement standpoint, the proposed measure also seeks to minimize potential confounds with working memory and attention to detail. While the concept of the measure is grounded in existing assessment approaches, its application is novel and may take some experimentation with various types of item stimuli and response sets. Given the novel nature of the Perceptive construct outside the realm of social perception, it would be useful to engage in further research to understand the nomological network surrounding the construct.

Interdisciplinary and Technologically Adept

While both Interdisciplinary and Technologically Adept are potentially important cognitive constructs for officers, our initial research findings suggested that additional conceptual work is needed before developing an assessment of these two constructs. The primary hurdle for Interdisciplinary is the need to determine what would be considered appropriate "disciplines" for military officers; whether it should be general fields of study such as education, psychology, business, etc., or something oriented to the military such as various branches or warfighter functions. A related question is determining at what level it would be appropriate to assess the construct, given the nature of the "disciplines" that are selected for the construct. In order to "synthesize and apply knowledge" from the various discipline, they need to be assessed in an officer who would be expected to have knowledge of those disciplines. It may be that it is not appropriate to assess this construct in newly commissioned officers, but rather is something that would need to be tested at the field grade officer level.

Measuring the Technologically Adept construct was expected to be challenging for multiple reasons. The construct was defined as "learns how to use and apply advances in technologies or technological systems"; however, one operating assumption was that the measure should test the skill directly rather than measure self-report of behaviors. In order to develop a skill-based measure we would need to assess participants as they were learning and using advances in technologies; however, another operating assumption was that the assessment should not require pre-existing knowledge. Assessing an officer directly on learning and using advances in technologies that already exist, this would give an advantage to individuals who already have a strong baseline knowledge of those technologies. One way around this predicament would be to invent fictitious technologies to prevent people from having an advantage. It will take a significant amount of time to enable test takers to learn about a fictitious technology and then apply it, which is also a problem, given our operating assumption that the measures should be relatively short. There was no assessment solution that fits within all of these parameters. While an assessment of Technologically Adept may be useful for assessing officers, it is not feasible within this battery of cognitive skills. Rather, technological requirements of officers should be identified, and more tailored approaches taken.

Priorities for Development

Each of the six assessment concepts recommended for development in this research would potentially be valuable elements of a cognitive battery for officers. While all are recommended for further development, there are four factors that could be considered as guidelines to prioritize constructs for development. The first factor that could be considered is their importance across all branches. Of the six constructs recommended for further development, several of these received mean ratings of "Very important" or "Extremely important" across all or nearly all branches (U.S. Army Research Institute, 2019). Analytical thinking from the ARI cognitive constructs, and detail focused, Problem Solver, and multi-tasker from the TAB received mean ratings of 3.00 (Very important) or higher across all branches. Logical/analytical from the TAB received a mean rating of 2.75 for one branch, but the other means were all above 3.00. Given the high ratings of importance across branches, these constructs are strong candidates to use in identifying high performing officers. The ability to Analyze Data or Information is important in the day-to-day responsibilities of officers as they are

constantly confronted with a variety of information that they need to understand and break down in order to apply, take appropriate action with, and use to solve problems.

Another factor that could be considered in determining priorities is identifying constructs that are most likely to distinguish between branches, thus facilitating use of the assessment for officer classification or reclassification purposes. Two constructs from the ARI construct list, process information and data (which was renamed Detailed Information Processing) and systems thinking, showed a split in their importance data, such that for a portion of the branches, the mean importance of the construct was above 3.00 and for the other branches the mean importance was below 3.00 (U.S. Army Research Institute, 2019). If these two constructs were included in a cognitive battery for officers, they could potentially be useful in determining which officers would be better suited for success in specific branches.

A third factor that could be considered is the extent to which the constructs have conceptual clarity, and their proposed measures have empirical support in the existing literature or in practice. The three constructs recommended for development that had the strongest conceptual and empirical support for their proposed measures were Analytical Thinking, Multitasking, and Detailed Information Processing. This suggests that choosing to develop these constructs would be somewhat less risky and the measures could potentially be ready for operational use with fewer iterations of development and testing.

Conversely, a final factor that could be considered is the extent to which the measure proposes an innovative or novel approach that would provide significant benefit if successful but is risky due to the untested nature of the measurement approach. While novel approaches hold more risk and may take longer to develop, that risk can result in significant future benefits. Three constructs that are based on a firm conceptual foundation but are novel in their approach to measuring a given construct are Perceptive, Problem Solver, and systems thinking. Each of these has potential to succeed but uses a novel approach for the construct so may take longer to be ready for operational use.

Skill-based vs Self-report Measures

One of the operating assumptions of this research was that measurement strategies should focus on measuring the cognitive skill or ability directly rather than through self-reported behaviors. Though a standard self-report approach is not preferred, there could be benefits to using a forced-choice self-report framework to measure a number of these constructs. A multi-dimensional forced-choice approach could be used for measuring a variety of constructs with one administration format. If properly calibrated with statements of equivalent social desirability this approach can reduce faking. Moreover, multi-dimensional forced-choice can be used with computer adaptive testing, reducing the number of items a test-taker needs to complete and shortening testing time.

Conclusions

Based on the four prioritization factors that were described, one of the constructs that should be prioritized for development is Analytical Thinking, which was identified as very important across all branches and has a high probability for successfully predicting officer

performance, given that similar measures have already been developed and validated. One activity that continues to be important is work by the Army to organize the various attributes into broader talent areas and identify and reduce overlap among the constructs. Another factor to consider in the talent framework is that, while the constructs identified may be conceptually distinct, they may be extremely difficult to distinguish operationally. Examples of this are the constructs Analytical Thinking and Analyze Data or Information, which are both in the Reasoning talent of the ATAF. While a conceptual distinction was made in the definitions of these two constructs, from a measurement standpoint it was not feasible to capture the conceptual distinctions. Thus, while ATAF provides a useful comprehensive list of relevant constructs, it may be useful to combine constructs that cannot be operationally distinguished or use the higher level "talent" categories as a tool to identify constructs that can be distinguished operationally.

Key cognitive abilities are required for officer success in making decisions, carrying out mission command, and working and adapting within the dynamic military environment (ADP 6-22, Army Leadership and the Profession). Various cognitive abilities have been identified as important for specific branches as well as across branches. Improving our assessment of officers across these cognitive abilities and skills has the potential to provide valuable feedback regarding these important capabilities and to provide important information to the Army regarding the classification of new officers into the various branches.

References

Adis, C., Wind, A. P., Wisecarver, M., Byrd, C., Martin, J., Engelsted, L., & Canali, K. G. (2022). *Common cyber capabilities assessment*. (Technical Report 1409). U.S. Army Research Institute for the Behavioral and Social Sciences.

Adis, C., Wisecarver, M., Hartzler, C., Mangos, P., Hulse, N., Wind, A., & Canali, K. G. (2016). *Proof of concept: Systems thinking assessment game*. (Contract Report). U.S. Army Research Institute for the Behavioral and Social Sciences.

Adis, C., Wisecarver, M., Byrd, C., Wind, A., & Canali, K. G. (2017). *Innovative tools to assess systems thinking ability*. (Technical Report 1362). U.S. Army Research Institute for the Behavioral and Social Sciences.

Alefari, M., Barahona, A. M. F., & Salonitis, K. (2018). Modelling manufacturing employees' performance based on a system dynamics approach. *Procedia CIRP*, *72*, 438-443. https://doi.org/10.1016/j.procir.2018.03.161

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/10.1111/j.1467-6486.1996.tb00801.x

Ambady, N., & Rosenthal, R. (1992). Thin slices of expressive behavior as predictors of interpersonal consequences: A meta-analysis. *Psychological Bulletin*, *111*(2), 256-274. https://doi.org/10.1037/0033-2909.111.2.256

Ambady, N., Bernieri, F. J., & Richeson, J. A. (2000). Toward a histology of social behavior: Judgmental accuracy from thin slices of the behavioral stream. In M. P Zanna (Ed.) *Advances in experimental social psychology* (Vol. 32, pp. 201-271). Academic Press. https://doi.org/10.1016/S0065-2601(00)80006-4

Arnold, R. D., & Wade, J. P. (2017). A complete set of systems thinking skills. *Insight*, *20*(3), 9–17. https://doi.org/10.1002/inst.12159.

Austen, E., & Enns, J. T. (2000). Change detection: Paying attention to detail. *Psyche: An Interdisciplinary Journal of Research on Consciousness*, *6*(11), 808-817. http://psyche.cs.monash.edu.au/v6/psyche-6-11-austen.html

Baird, L. L. (1983). Review of problem-solving skills. *ETS Research Report Series*, *1983*(1), i-45. https://doi.org/10.1002/j.2330-8516.1983.tb00016.x

Baksh, R. A., Abrahams, S., Auyeung, B., & MacPherson, S. E. (2018). The Edinburgh Social Cognition Test (ESCoT): Examining the effects of age on a new measure of theory of mind and social norm understanding. *PLoS ONE, 13*(4), Article e0195818. https://doi.org/10.1371/journal.pone.0195818

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review, 84*(2), 191–215. https://doi.org/10.1037/0033-295X.84.2.191

Arca, V., Hurrell, A., Macauslan, I., Visram, A., & Willis, J. (2013). Paying attention to detail: how to transfer cash in cash transfers. *Enterprise Development & Microfinance, 24*(1), 10–27. https://doi.org/10.3362/1755-1986.2013.003

Baylis, G. C. (1994). Visual attention and objects: Two-object cost with equal convexity. *Journal of Experimental Psychology: Human Perception and Performance*, *20*(1), 208. https://doi.org/10.1037/0096-1523.20.1.208

Baylis, G. C., & Driver, J. (1992). Visual parsing and response competition: The effect of grouping factors. *Perception & Psychophysics*, *51*(2), 145-162. https://doi.org/10.3758/BF03212239

Beck A. T., Baruch E., Balter J. M., Steer R. A., & Warman D. M. (2004). A new instrument for measuring insight: The Beck Cognitive Insight Scale. *Schizophrenia Research*, *68*(2/3), 319-329. https://doi.org/10.1016/S0920-9964(03)00189-0

Bell, M. S. (1979). Applied problem solving as a school emphasis: An assessment and some recommendations (ED180816). In R. Lesh, D. Mierkiewicz, and M. Kantowski, (Eds.), *Applied mathematical problem solving* (pp. 5-24). ERIC Clearinghouse for Science.

Booth Sweeney, L., & Sterman, J. (2000). Bathtub dynamics: Initial results of a systems thinking inventory. *System Dynamics Review, 16*(4), 249–286. https://doi.org/10.1002/sdr.366

Borg, J. R., & Borg, M. O. (2001). Teaching critical thinking in interdisciplinary economics courses. *College Teaching*, *49*(1), 20-25. https://doi.org/10.1080/87567550109595841

Boyatzis, R. E., & Goleman, D. (2007). *Emotional and social competency inventory - University edition workbook*. Hay Group Transforming Learning.

Boyatzis, R. E., & Kolb, D. A. (1995). From learning styles to learning skills: The executive skills profile. *Journal of Managerial Psychology, 10*(5), 3-17. https://doi.org/10.1108/02683949510085938

Boyle, B. E., Darbor, K. E., Cato, C. R., Kammerer, W., & Thurston, J. (2022). *Survey of Talent Requirements: Talents, Skills, and Abilities of Army Officers by Branch and Functional Area* (Technical Report 1418). U.S. Army Research Institute for the Behavioral and Social Sciences.

Bransford, J. D., & Stein, B. S. (1993). *The IDEAL problem solver: A guide for improving thinking, learning, and creativity.* W. H. Freeman.

Brod, C. (1984). *Technostress: The human cost of the computer revolution*. Addison-Wesley.

Brubacher, S. P., Gilligan, C., Burrows, K. S., & Powell, M. B. (2019). Information gathering in investigative and medical interviewing: Drawing parallels across contexts. *Health Communication*, *36*(4), 521-528. https://doi.org/10.1080/10410236.2019.1700884

Burgess, P. W., Veitch, E., de Lacy Costello, A., & Shallice, T. (2000). The cognitive and neuroanatomical correlates of multitasking. *Neuropsychologia*, *38*(6), 848-863. https://doi.org/10.1016/S0028-3932(99)00134-7

Burns, J. C., Okey, J. R., & Wise, K. C. (1985). Development of an integrated process skill test: TIPS II. *Journal of Research in Science Teaching*, *22*(2), 169-177. https://doi.org/10.1002/tea.3660220208

Butler, H. A., Dwyer, C. P., Hogan, M. J., Franco, A., Rivas, S. F., Saiz, C., Almeida, L. S. (2012). The Halpern critical thinking assessment and real-world outcomes: Cross-national applications. *Thinking Skills and Creativity, 7(2),* 112-121. https://doi.org/10.1002/acp.2851

Cappel, J. J., Prybutok, V. R., & Varghese, B. (2005). A closer look at attention to detail. *Communications of the ACM*, *48*(7), 87-92. https://doi.org/10.1145/1070838.1070842

Carrier, L. M., Cheever, N. A., Rosen, L. D., Benitez, S., & Chang, J. (2009). Multitasking across generations: Multitasking choices and difficulty ratings in three generations of Americans. *Computers in Human Behavior*, *25*(2), 483-489. https://doi.org/10.1016/j.chb.2008.10.012

Chaiken, S. R. (1994). The inspection time not studied: Processing speed ability unrelated to psychometric intelligence. *Intelligence, 19*(3), 295-313. https://doi.org/10.1016/0160-2896(94)90003-5

Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology, 4*(1)*,* 55-81. https://doi.org/10.1016/0010-0285(73)90004-2

Chen, X., Hertzog, C., & Park, D. C. (2017). Cognitive predictors of everyday problem solving across the lifespan. *Gerontology*, *63*(4), 372-384. https://doi.org/10.1159/000459622

Chi, M. T. H., Glaser, R., & Rees, E. (1982). Expertise in problem solving. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (Vol 1, pp.7-75.). Lawrence Erlbaum Associates, Inc.

Chiang, R. H., Goes, P., & Stohr, E. A. (2012). Business intelligence and analytics education, and program development: A unique opportunity for the information systems discipline. *ACM Transactions on Management Information Systems (TMIS), 3*(3), 1-13. https://doi.org/10.1145/2361256.2361257

Colberg, M., Nester, M. A., & Trattner, M. H. (1985). Convergence of the inductive and deductive models in the measurement of reasoning abilities. *Journal of Applied Psychology, 70*(4), 681. https://doi.org/10.1037/0021-9010.70.4.681

Colom, R., Martínez-Molina, A., Shih, P. C., & Santacreu, J. (2010). Intelligence, working memory, and multitasking performance. *Intelligence*, *38*(6), 543-551. https://doi.org/10.1016/j.intell.2010.08.002

Compeau, D. R., and Higgins, C. A. (1995). Computer Self-Efficacy: Development of a Measure and Initial Test, *MIS Quarterly, 19*(2), 189-211. https://doi.org/10.2307/249688

Corte Real, N., Oliveira, T., & Ruivo, P. (2015, September 21-22). *Understanding the hidden value of business intelligence and analytics (BI&A)* [Paper presentation]. 9th European Conference on IS Management and Evaluation, The University of West England, UK. https://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.930.2463&rep=rep1&type=pdf

Cronin, J. J., Jr., Brady, M. K., & Hult, G. T. M. (2000). Assessing the effects of quality, value, and customer satisfaction on consumer behavioral intentions in service environments. *Journal of Retailing, 76*(2), 193–218. https://doi.org/10.1016/S0022-4359(00)00028-2

Cyert, R. M. (1980). Problem solving and educational policy. In D. T. Tuma, and F. Reif, (Eds.). *Problem solving and education: Issues in teaching and research*. Lawrence Erlbaum Associates, Inc.

Davis, A. C., & Stroink, M. L. (2016). The Relationship between systems thinking and the new ecological paradigm. *System Research and Behavioral Science*, *33*(4), 575–586. https://doi.org/10.1002/sres.2371

Davis, F. D., Bagozzi, R. P., & Warshaw, P. R. (1989). User acceptance of computer technology: A comparison of two theoretical models. *Management Science, 35*(8), 982–1003. https://doi.org/10.1287/mnsc.35.8.982

Davis, G. A. (1973). *Psychology of problem solving: theory and practice*. Basic Books.

Davis. G. A. (1966). Current status of research and theory in human problem solving. *Psychological Bulletin*, *66*(1), 36-54. https://doi.org/10.1037/h0023460

de Groot, A. (1966). Perception and memory versus thought: Some old ideas and recent findings. In B. Kleinmuntz (Ed.), *Problem Solving: Research, Method and Theory* (pp.19-50). John Wiley & Sons.

de Queiroz Tourinho, A. L., Sanchez, O. P., & Brown, S. A. (2019, June 8-14). *Measuring the organizational analytical competence: Development of a scale* [Paper presentation]. 27th European Conference on Information Systems (ECIS), Stockholm and Uppsala, Sweden. https://core.ac.uk/download/pdf/301379442.pdf

Delbridge, K. A. (2000). *Individual differences in multi-tasking ability: Exploring a nomological network*. [Doctoral dissertation, Michigan State University]. Michigan State University Digital Repository. https://doi.org/doi:10.25335/M5WW77B5J

Eastin, M. S., & LaRose, R. (2000). Internet self-efficacy and the psychology of the digital divide. *Journal of Computer-Mediated Communication, 6*(1), Article JCMC611. https://doi.org/10.1111/j.1083-6101.2000.tb00110.x

Engel, G. L. (1977). The need for a new medical model: a challenge for biomedicine. *Science*, *196*(4286), 129-136. https://doi.org/10.3109/13561828909043606

Engelmann, A., Kump, B., & Schweiger, C. (2020). Clarifying the Dominant Logic Construct by Disentangling and Reassembling its Dimensions. *International Journal of Management Reviews*, *22*(4), 323-355. https://doi.org/10.1111/ijmr.12227

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*(2), 390-405. https://doi.org/10.1037/0022-3514.71.2.390

Eriksen, C. W., & Shultz, D. W. (1979). Information processing in visual search: A continuous flow conception and experimental results. *Perception & Psychophysics, 25*(4), 249-263. https://doi.org/10.3758/bf03198804

Faust, M. E., Balota, D. A., Spieler, D. H., & Ferraro, F. R. (1999). Individual differences in information-processing rate and amount: Implication for group differences in response latency. *Psychological Bulletin, 125*(6), 777-799. https://doi.org/10.1037/0033-2909.125.6.777

Ferrari, A. (2012). *Digital competence in practice: An analysis of frameworks*. JRC IPTS.

Fischer, H. R. (2001). Abductive reasoning as a way of worldmaking. *Foundations of Science, 6*(4), 361-383. https://www.univie.ac.at/constructivism/pub/fos/pdf/fischer.pdf

Forrester, J. F. (1961). *Industrial dynamics*. Pegasus Communications. https://doi.org/10.1057/palgrave.jors.2600946

Frideres, J.S., Goldenberg, S., Disanto, J., & Fleming, U. (1983). Incidence and potential causal factors. *Social Indicators Research, 13*(4), 381–393. https://doi.org/10.1007/bf00318077

Funke, J. (2010). Complex problem solving: A case for complex cognition?. *Cognitive Processing*, *11*(2), 133-142. https://doi.org/10.1007/s10339-009-0345-0

Gantt, L. O., & Natt, I. I. (2006). Deconstructing thinking like a lawyer: Analyzing the cognitive components of the analytical mind. *Campbell Law Review, 29*(3), 413-481. https://doi.org/10.2139/ssrn.1953470

Ghasemaghaei, M. (2019). Does data analytics use improve firm decision making quality? The role of knowledge sharing and data analytics competency. *Decision Support Systems, 120*, 14-24. https://doi.org/10.1016/j.dss.2019.03.004

Ghasemaghaei, M., Ebrahimi, S., & Hassanein, K. (2018). Data analytics competency for improving firm decision making performance. *The Journal of Strategic Information Systems, 27*(1), 101-113. https://doi.org/10.1016/j.jsis.2017.10.001

Giacomini C.L., Lyle H.III, Wynn W. (2012). Developing a Context-Specific Measure of Technological Expertise: The University of Washington. *Educause Seeking Evidence of Impact.* http://www.washington.edu/lst/research/research_projects/2011techsurveys

Gil, Y. (2016, February). *Teaching Big Data Analytics Skills with Intelligent Workflow Systems* [Paper presentation]. National Conference of the Association for the Advancement of Artificial Intelligence, Phoenix, Arizona, United States. https://ojs.aaai.org/index.php/AAAI/article/view/9860

Girelli, L., Semenza, C., & Delazer, M. (2004). Inductive reasoning and implicit memory: evidence from intact and impaired memory systems. *Neuropsychologia, 42*(7), 926-938. https://doi.org/10.1016/j.neuropsychologia.2003.11.016

Glass, A. L., & Holyoak, K. J. (1986). *Cognition* (2nd ed.). Random House.

Goodhue, D. L. (1998). Understanding user evaluations and information systems. *Management Science, 41*(12), 1827-1844. https://doi.org/10.1287/mnsc.41.12.1827

Goodhue, D. L., & Thompson, R. L. (1995). Task-technology fit and individual performance. *MIS Quarterly, 19*(2), 213-236. https://doi.org/10.2307/249689

Goodwin, G. F., Lockhart, G., Keil, C., Brewington, H., Ness, A., & Bolger, K. (2020, August 19). *Army talent attribute framework.* [Unpublished briefing]. U.S. Army Research Institute for the Behavioral and Social Sciences.

Gori, A., Craparo, G., Giannini, M., Loscalzo, Y., & Caretti, V. (2015). Development of a new measure for assessing insight: Psychometric properties of the insight orientation scale (IOS). *Schizophrenia Research*, *169*(1/3), 298-302. https://doi.org/10.1016/j.schres.2015.10.014

Gottesman, L., & Chapman, L. J. (1960). Syllogistic reasoning errors in schizophrenia. *Journal of Consulting Psychology*, *24*(3), 250-255. https://doi.org/10.1037/h0042189

Gottsdanker, R., & Schragg, G. P. (1985). Verification of Donders' subtraction method. *Journal of Experimental Psychology: Human Perception and Performance, 11*(6), 765-776. https://doi.org/10.1037/0096-1523.11.6.765

Greene, J. A., Azevedo, R., & Torney-Purta, J. (2008). Modeling epistemic and ontological cognition: Philosophical perspectives and methodological directions. *Educational Psychologist, 43*(3), 142–160. https://doi.org/10.1080/00461520802178458.

Greeno, J. G. (1989). A perspective on thinking. *American Psychologist, 44*(2), 134-141. https://doi.org/10.1037/0003-066X.44.2.134

Griffiths, A. K., Grant, B. A. (1985). High school students' understandings of food webs: Identification of a learning hierarchy and related misconceptions. *Journal of Research in Science Teaching, 22*(5), 421–436. https://doi.org/10.1002/tea.3660220505

Grimes, J. (1996). On the failure to detect changes in scenes across saccades. In K. Akins, *Perception* (pp. 89-110). Oxford University Press.

Grohs, J. R., Kirk, G. R., Soledad, M. M., & Knight, D. B. (2018). Assessing systems thinking: A tool to measure complex reasoning through ill-structured problems. *Thinking Skills and Creativity*, *28*, 110-130. https://doi.org/10.1016/j.tsc.2018.03.003

Hall, K. L., Stokols, D., Moser, R. P., Taylor, B. K., Thornquist, M. D., Nebeling, L. C., & Jeffery, R. W. (2008). The collaboration readiness of transdisciplinary research teams and centers: Findings from the National Cancer Institute's TREC year-one evaluation study. *American Journal of Preventive Medicine*, *35*(2), S161-S172. https://doi.org/10.1016/j.amepre.2008.03.035

Hambrick, D. Z., Oswald, F. L., Darowski, E. S., Rench, T. A., & Brou, R. (2010). Predictors of multitasking performance in a synthetic work paradigm. *Applied Cognitive Psychology*, *24*(8), 1149-1167. https://doi.org/10.1002/acp.1624

Happé, F., & Frith, U. (2006). The weak coherence account: Detail-focused cognitive style in autism spectrum disorders. *Journal of Autism and Developmental Disorders*, *36*(1), 5-25. https://doi.org/10.1007/s10803-005-0039-0

Hargittai E. (2005) Survey measures of web-oriented digital literacy. *Social Science Computer Review, 23*(3), 371-379. https://doi.org/10.1177/0894439305275911

Harvey, O. J., Hunt, D. E., & Schroder, H. M. (1961). *Conceptual systems and personality organization.* John Wiley & Sons.

Hättenschwiler, N., Merks, S., & Schwaninger, A. (2018, October). *Airport security X-ray screening of hold baggage: 2D versus 3D imaging and evaluation of an on-screen alarm resolution protocol* [Paper presentation]. 2018 International Carnahan Conference on Security Technology, Montreal, Canada. https://doi.org/10.1109/CCST.2018.8585713

Hattingh, M., Marshall, L., Holmner, M., & Naidoo, R. (2019, September 17-18). *Data science competency in organisations: A systematic review and unified model* [Paper presentation]. The South African Institute of Computer Scientists and Information Technologists, Skukuza, South Africa. https://doi.org/10.1145/3351108.3351110

Hayes, J.R. (1989). Cognitive Processes in Creativity. In Glover, J.A., Ronning, R.R., Reynolds, C.R. (eds) *Handbook of Creativity. Perspectives on Individual Differences* (135-145). Springer. https://doi.org/10.1007/978-1-4757-5356-1_7

Heinssen, R.K., Glass, C.R., & Knight, L.A. (1987). Assessing computer anxiety: development and validation of the computer anxiety rating scale. *Computers in Human Behavior, 3*(1) 49–59. https://doi.org/10.1016/0747-5632(87)90010-0

Hennessy, C. H., & Walker, A. (2011). Promoting multi-disciplinary and inter-disciplinary ageing research in the United Kingdom. *Ageing & Society*, *31*(1), 52-69. https://doi.org/10.1017/S0144686X1000067X

Holzman, T., Pellegrino, J., & Glaser, R., (1983). Cognitive variables in series completion. *Journal of Educational Psychology, 75*(4), 603–618. https://doi.org/10.1037/0022-0663.75.4.603

International Integrated Reporting Council. (2013). *The International Integrated Reporting Framework.* International Integrated Reporting Council. https://integratedreporting.org/resource/international-ir-framework/

Jafar, M. J., Babb, J. S., & Abdullat, A. (2017). Emergence of data analytics in the information systems curriculum. *Information Systems Education Journal, 15*(5), 22-36. http://isedj.org/2017-15/n5/ISEDJv15n5p22.html

Jay, T. B. (1981). Computerphobia: What to do about it? *Educational Technology, 21,* 47–48.

Johannes, M. N., & Kasteren, V. (1996). Interdisciplinary teaching within engineering education. *European Journal of Engineering Education*, *21*(4), 387-392. https://doi.org/10.1080/03043799608923424

Johnson-Laird, P. N. (1995). Mental models, deductive reasoning, and the brain. *The Cognitive Neurosciences* (pp. 999-1008). https://doi.org/10.1017/cbo9780511814273.012

Joyce, M., Kirakowski, J. (2014). Measuring Confidence in Internet Use: The Development of an Internet Self-efficacy Scale. In Marcus, A. (Eds.), *Design, User Experience, and Usability. Theories, Methods, and Tools for Designing the User Experience* (pp. 250-260). Springer. https://doi.org/10.1007/978-3-319-07668-3_25

Kandalaft, M. R., Didehbani, N., Cullum, C. M., Krawczyk, D. C., Allen, T. T., Tamminga, C. A., & Chapman, S. B. (2012). The Wechsler ACS social perception subtest a preliminary comparison with other measures of social cognition. *Journal of Psychoeducational Assessment, 30*(5), 455–465. https://doi.org/10.1177/0734282912436411

Kingry, D., Engelsted, J., Van DeGriek, O., Gray, C., Zorzie, M., Madaj, C. A., Bruskiewicz, K. T., & Bobrow, W. (2020). *Development and validation of an assessment battery for the FBI intelligence analyst position*. (Technical Report No. 996). Personnel Decisions Research Institutes, LLC.

Kirton, M. (1976). Adaptors and innovators: A description and measure. *Journal of Applied Psychology*, *61*(5), 622-629. https://doi.org/10.1037/0021-9010.61.5.622

Kitchen, P., & Spickett-Jones, G. (2003). Information processing: A critical literature review and future research directions. *International Journal of Market Research, 45*(1), 73-98. https://doi.org/10.1177/147078530304500106

Klein, J., & Newell, W. (1997). Advancing interdisciplinary studies. In J. Gaff and J. Ratcliffe (Eds.), *Handbook of the undergraduate curriculum: A comprehensive guide to purposes, structures, practices, and changes* (pp. 393-415). Jossey-Bass.

Klein, J. (1999). The relationship between level of academic education and reversible and irreversible processes of probability decision-making. *Higher Education, 37*(4), 323-339. https://doi.org/10.1023/A:1003608412240

Klein, J. T. (1990). *Interdisciplinarity: History, theory, and practice*. Wayne State University Press.

Klein, J. T. (1999). *Mapping Interdisciplinary Studies. The Academy in Transition*. Association of American Colleges and Universities.

Klix, F. (2001). Problem solving: Deduction, induction, and analogical reasoning. In N. J. Smelser and P. B. Baltes (Eds.), *International Encyclopedia of the Social & Behavioral Sciences* (pp.12123-12130). Pergamon.

Konig, C. J., Buhner, M., & Murling, G. (2005). Working memory, fluid intelligence, and attention are predictors of multitasking performance, but polychronicity and extraversion are not. *Human Performance*, *18*(3), 243-266. https://doi.org/10.1207/s15327043hup1803_3

Korukonda, A. R. (2005). Personality, individual characteristics, and predisposition to technophobia: Some answers, questions, and points to ponder about. *Information Sciences, 170*(2/4), 309-328. https://doi.org/10.1016/j.ins.2004.03.007

Kotovsky, K., & Simon, H. A. (1973). Empirical tests of a theory of human acquisition of concepts for sequential patterns. *Cognitive Psychology*, *4*(3), 399-424. https://doi.org/10.1016/0010-0285(73)90020-0

Kuhn, D. (1977). Conditional reasoning in children. *Developmental Psychology, 13*(4), 342. https://doi.org/10.1037/0012-1649.13.4.342

Kvedaravicius, J. E., Skarzauskiene, A., & Palaima, T. (2009). Systems thinking as intelligence competence and its relationship to leadership performance. *Problems and Perspectives in Management*, *7*(2), 75-85. https://hdl.handle.net/20.500.12259/53220

Kwon, Y. J., Lee, J. K., Shin, D. H., & Jeong, J. S. (2009). Changes in brain activation induced by the training of hypothesis generation skills: An fMRI study. *Brain and Cognition*, *69*(2), 391-397. https://doi.org/10.1016/j.bandc.2008.08.032

Kyllonen, P. C. (1985). Dimensions of information processing speed. (AFHRL-TP-84-56). Air Force Human Resources Lab. https://apps.dtic.mil/sti/citations/ADA154778

LaFrance, D. L., Weiss, M. J., Kazemi, E., Gerenser, J., & Dobres, J. (2019). Multidisciplinary teaming: Enhancing collaboration through increased understanding. *Behavior Analysis in Practice*, *12*(3), 709-726. https://doi.org/10.1007/s40617-019-00331-y

Larkin, J. H. (1980). Teaching problem solving in physics: The psychological laboratory and the practical classroom. In D. T. Tuma & F. Reif, (Eds.), *Problem solving and education: Issues in teaching and research* (pp.111-125). Lawrence Erlbaum Associates.

Lee, H., Chang, H., Choi, K., Kim, S. W., & Zeidler, D. L. (2012). Developing character and values for global citizens: Analysis of pre-service science teachers' moral reasoning on socioscientific issues. *International Journal of Science Education*, *34*(6), 925-953. https://doi.org/10.1080/09500693.2011.625505

Lefevre, J., & Bisanz, J., (1986). A cognitive analysis of number series problems: Sources of individual differences in performance. *Memory & Cognition, 14*(4), 287–298. https://doi.org/10.3758/BF03202506

Levin, D. T., & Simons, D. J. (1997). Failure to detect changes to attended objects in motion pictures. *Psychonomic Bulletin & Review*, *4*(4), 501-506. https://doi.org/10.3758/BF03214339

Lohman, D. F. (1989). Estimating individual differences in information processing using speed-accuracy models. In R. Kanfer, P. L. Ackerman, & R. Cudeck (Eds.), *Abilities, Motivation and Methodology*: *The Minnesota Symposium on Learning and Individual Differences* (pp. 119–163). Routledge.

Lord, R. G. (1985). An information-processing approach to social perceptions, leadership and behavioral measurement in organizations. *Research in Organizational Behavior*, *7*, 87-128. https://psycnet.apa.org/record/1986-02663-001

Lord, R.G. & K.J. Maher (1990). Alternative information-processing models and their implications for theory, research, and practice. *The Academy of Management Review, 15*(1), 9-28. https://doi.org/10.5465/amr.1990.4308219

Lucas, H. D., Duff, M. C., & Cohen, N. J. (2019). The hippocampus promotes effective saccadic information gathering in humans. *Journal of Cognitive Neuroscience*, *31*(2), 1–16. https://doi.org/10.1162/jocn_a_01336

Lyons, N. R. (1986). *The basis for technological literacy: Experiments in exploration*. (Technical Report #1986-08). Naval Postgraduate School. https://doi.org/10.1.1.1004.4429

Mansilla, V. B., Dillon, D., & Middlebrooks, K. (2000). *Building bridges across disciplines: organizational and individual qualities of exemplary interdisciplinary work. Interdisciplinary studies project* (Paper 16: GoodWork Project, Project Zero). Harvard Graduate School of Education.

Mayer, J. D.; Salovey, P., & Caruso, D. R. (2002) *Mayer-Salovey-Caruso Emotional Intelligence Test (MSCEIT) Item Booklet*. Multi-Health Systems Assessments. https://scholars.unh.edu/personality_lab/26

McDonald, S., Fisher, A., Togher, L., Tate, R., Rushby, J., English, T., Kelly, M., Mathersul, D., Froreich, F., & Francis, H. (2015). Adolescent performance on The Awareness of Social Inference Test: TASIT. *Brain Impairment*, *16*(1), 3-18. https://doi.org/10.1017/BrImp.2015.7

McDonald, T., & Siegall, M. (1992). The Effects of Technological Self-Efficacy and Job Focus on Job Performance, Attitudes, and Withdrawal Behaviors. *The Journal of Psychology Interdisciplinary and Applied 126(5),* 465-475. https://doi.org/10.1080/00223980.1992.10543380

McLean, G. W. (2012). *Leadership principles for the new ADP 6-22*. Army Command and General Staff College, School of Advanced Military Studies. https://apps.dtic.mil/sti/citations/ADA566527

Meier, S. T. (1985) Computer aversion. *Computers in Human Behavior, 1*(2), 171–179. https://doi.org/10.1016/0747-5632(85)90030-5

Meissner, C., Redlich, A., Michael, S., Evans, J., Camilletti, C., Bhatt, S., & Brandon, S. (2014). Accusatorial and information-gathering interrogation methods and their effects on true and false confessions: A meta-analytic review. *Journal of Experimental Criminology*, *10*(4), 459–486. https://doi.org/10.1007/s11292-014-9207-6

Mellouli, N., & Bouchon-Meunier, B. (2003). Abductive reasoning and measures of similitude in the presence of fuzzy rules. *Fuzzy Sets and Systems, 137*(1), 177-188. https://doi.org/10.1016/S0165-0114(02)00439-6

Mendenhall, T. J., & Berge, J. M. (2010). Family therapists in trauma-response teams: Bringing systems thinking into interdisciplinary fieldwork. *Journal of Family Therapy, 32*(1), 43-57. https://doi.org/10.1111/j.1467-6427.2009.00482.x

Merks, S., Hättenschwiler, N., Zeballos, M., & Schwaninger, A. (2018, October). *X-ray screening of hold baggage: Are the same visual-cognitive abilities needed for 2D and 3D imaging?* [Paper presentation]. 2018 International Carnahan Conference on Security Technology. https://doi.org/10.1109/CCST.2018.8585715

Merriam-Webster. (n.d.) Compile, code, categorize, tabulate, audit, verify. In Merriam-Webster.com dictionary. Retrieved February 19, 2022, from https://www.merriam-webster.com/dictionary/

Merritt K., Smith D., & Renzo J. C. D. (2005). An investigation of self-reported computer literacy: Is it reliable? *Issues in Information Systems, 6*(1), 289-295. https://doi.org/10.48009/1_iis_2005_289-295

Meyer, M. S., & Cranmore, J. (2020). Supporting the information gathering practices of academically advanced college-bound students. *Psychology in the Schools*, *57*(7), 1057–1076. https://doi.org/10.1002/pits.22372

Miller, J. C., Redlich, A. D., & Kelly, C. E. (2018). Accusatorial and information-gathering interview and interrogation methods: A multi-country comparison. *Psychology, Crime & Law*, *24*(9), 935–956. https://doi.org/10.1080/1068316X.2018.1467909

Miron-Spektor, E., Erez, M., & Naveh, E. (2011). The effect of conformist and attentive-to-detail members on team innovation: Reconciling the innovation paradox. *Academy of Management Journal*, *54*(4), 740-760. https://doi.org/10.5465/amj.2011.64870100

Miron, E., Erez, M., & Naveh, E. (2004). Do personal characteristics and cultural values that promote innovation, quality, and efficiency compete or complement each other?. *Journal of Organizational Behavior*, *25*(2), 175-199. https://doi.org/10.1002/job.237

Mohaghegh, M., & Furlan, A. (2020). Systematic problem-solving and its antecedents: a synthesis of the literature. *Management Research Review, 43*(9), 1033-3062. https://doi.org/10.1108/MRR-06-2019-0284

Moore, S. M., Dolansky, M. A., Singh, M., Palmieri, P., & Alemi, F. (2010). *The Systems Thinking Scale*. Case Western Reserve University. https://case.edu/nursing/sites/case.edu.nursing/files/2018-04/STS_Manual.pdf

Muraki, E. (1992). A generalized partial credit model: Application of an EM algorithm. *Applied Psychological Measurement, 16*(2)*,* 159-176. https://doi.org/10.1002/j.2333-8504.1992.tb01436.x

Muraki, E. (1993). Information functions of the generalized partial credit model. *Applied Psychological Measurement, 17*(4), 351-363. https://doi.org/10.1002/j.2333-8504.1993.tb01538.x

Nandiwada, D. R., & Dang-Vu, C. (2010). Transdisciplinary health care education: Training team players. *Journal of Health Care for the Poor and Underserved*, *21*(1), 26-34. http://doi.org/10.1353/hpu.0.0233

Naveh, E., & Erez, M. (2004). Innovation and attention to detail in the quality improvement paradigm. *Management Science*, *50*(11), 1576-1586. https://doi.org/10.1287/mnsc.1040.0272

Neubert, J. C. (2016). *Complex problem solving and organizational psychology: Of nomological networks, microworlds, and interdisciplinary dialogue*. [Unpublished doctoral dissertation]. Heidelberg University.

Newell, W. (1998). Professionalizing interdisciplinarity. In W. Newell (Ed.*), Interdisciplinarity: Essays from the literature* (pp. 529-563). College Board.

Newell, W. H. (2001). A theory of interdisciplinary studies. *Issues in Integrative Studies,* 19, 1-25. http://hdl.handle.net/10323/4378

Newell, A. (1980). Physical symbol systems. *Cognitive Science*, *4*(2), 135-183. https://doi.org/10.1016/S0364-0213(80)80015-2

Newell, A. & H.A. Simon (1972). *Human problem solving*. Prentice Hall.

Newell, W. H., & Green, W. J. (1982). Defining and teaching interdisciplinary studies. *Improving College and University Teaching*, *30*(1), 23-30. https://doi.org/10.1080/00193089.1982.10533747

Nijboer, M., Taatgen, N. A., Brands, A., Borst, J. P., & van Rijn, H. (2013). Decision making in concurrent multitasking: Do people adapt to task interference?. *PloS One*, *8*(11), Article e79583. https://doi.org/10.1371/journal.pone.0079583

Nilsson, N. J. (1971). *Problem-solving methods in artificial intelligence*. McGraw-Hill.

O'Reilly III, C. A., Chatman, J., & Caldwell, D. F. (1991). People and organizational culture: A profile comparison approach to assessing person-organization fit. *Academy of Management Journal*, *34*(3), 487-516. https://doi.org/10.5465/256404

O*NET Online (n.d.). Retrieved from http://www.onetonline.org/find/descriptor/browse/Skills/.

Özgelen, S. (2012). Students' science process skills within a cognitive domain framework. *Eurasia Journal of Mathematics, Science and Technology Education, 8*(4), 283-292. https://doi.org/10.12973/eurasia.2012.846a

Palaima, T., & Skarzauskiene, A. (2010). Systems thinking as a platform for leadership performance in a complex world. *Baltic Journal of Management, 5*(3), 330-355. https://doi.org/10.1108/17465261011079749

Pinkham, A. E., Penn, D. L., Green, M. F., & Harvey, P.D. (2016, March). Social cognition psychometric evaluation: Results of the initial psychometric study. *Schizophrenia Bulletin, 42*(2), 494–504, https://doi.org/10.1093/schbul/sbv056

Prahalad, C. K., & Bettis, R. A. (1986). The dominant logic: A new linkage between diversity and performance. *Strategic Management Journal, 7*(6), 485-501. https://doi.org/10.1002/smj.4250070602

Pretz, J. E., Naples, A. J., & Sternberg, R. J. (2003). Recognizing, defining, and representing problems. In J.E. Davidson & R.J. Sternberg (Eds.), *The Psychology of Problem Solving*. Cambridge University Press.

Raub, A. (1982). *Correlates of computer anxiety in college students*.(Publication No. 8208027). [Doctoral dissertation, University of Pennsylvania]. ProQuest Dissertations and Theses Global.

Rensink, R. A., O'Regan, J. K., & Clark, J. J. (1997). To see or not to see: The need for attention to perceive changes in scenes. *Psychological Science*, *8*(5), 368-373. https://doi.org/10.1111/j.1467-9280.1997.tb00427.x

Revlin, R. (2013). *Cognition theory and practice*. Worth Publishers.

Rieh, S. Y., & Belkin, N. J. (1998). Understanding judgment of information quality and cognitive authority in the WWW. *Journal of the American Society for Information Science and Technology, 35*, 279-89. https://www.researchgate.net/publication/238975063

Roberts, D. L., Fiszdon, J., & Tek, C. (2011). Initial validity of the Social Cognition Screening Questionnaire (SCSQ). *Schizophrenia Bulletin, 37*(Suppl 1), 280.

Rosen, L.D., & Weil, M.M. (1990a). Computers, classroom instruction, and the computer-phobic university student. *Collegiate Microcomputer, 8*(4), 275–283. https://doi.org/10.5555/98989.99000

Rosen, L.D., & Weil, M.M. (1990b). Myths and realities of computerphobia. *Anxiety Research, 3*(3)*,* 175–191. https://doi.org/10.1080/08917779008248751

Rosen, L.D., Sears, D.C., & Weil, M.M. (1987). Computerphobia. *Behavior Research Methods, Instrumentation, & Computers, 19(2),* 167–179. https://doi.org/10.3758/BF03203781

Rosen, L.D., Sears, D.C., & Weil, M.M. (1993). Treating technophobia: A longitudinal evaluation of the computerphobia reduction program. *Computers in Human Behavior, 9*(1)*,* 27–50.

Royston, R. P., & Lin, N. (In preparation). *Army-wide Job Analysis for Officers: Branch and Functional Area-specific Knowledge, Skills, and Behaviors Requirements from the Army Talent Attribute Framework: Main Report* (Technical Report). U.S. Army Research Institute for the Behavioral and Social Sciences.

Royston, R. P., & Lin, N. (In preparation). *Army-wide Job Analysis for Officers: Branch and Functional Area-specific Knowledge, Skills, and Behaviors Requirements from the Army Talent Attribute Framework: Appendices* (Research Note). U.S. Army Research Institute for the Behavioral and Social Sciences.

Rubenstein, M. F. (1975). *Patterns of problem solving*. Prentice Hall.

Ruddy, G., & Rhee, K. S. (2005). Transdisciplinary teams in primary care for the underserved: A literature review. *Journal of Health Care for the Poor and Underserved*, *16*(2), 248-256. https://doi.org/10.1353/hpu.2005.0040

Rusconi, E., McCrory, E., & Viding, E. (2012). Self-rated attention to detail predicts threat detection performance in security X-ray images. *Security Journal*, *25*(4), 356-371. https://doi.org/10.1057/sj.2011.27

Russom, P. (2011). Big data analytics. *TDWI Best Practices Report, Fourth Quarter*, *19*(4), 1-34. https://vivomente.com/wp-content/uploads/2016/04/big-data-analytics-white-paper.pdf

Salvucci, D. D., & Taatgen, N. A. (2008). Threaded cognition: An integrated theory of concurrent multitasking. *Psychological Review*, *115*(1), 101-130. https://doi.org/10.1037/0033-295X.115.1.101

Schoenfeld, A. H. (1979). Explicit heuristic training as a variable in problem-solving performance. *Journal for Research in Mathematics Education, 10(3),* 173-187. https://doi.org/10.5951/jresematheduc.10.3.0173

Schraw, G., Dunkle, M. E., & Bendixen, L. D. (1995). Cognitive processes in well-defined and ill-defined problem solving. *Applied Cognitive Psychology*, *9*(6), 523-538. https://doi.org/10.1002/acp.2350090605

Schroder, H. M., Driver, M.J. and Streufert, S. (1967). *Human information processing*. Holt, Rinehart and Winston.

Schvaneveldt, R.W., Cohen, T.A. (2010). Abductive Reasoning and Similarity: Some Computational Tools. In Ifenthaler, D., Pirnay-Dummer, P., Seel, N. (Edst.), *Computer-Based Diagnostics and Systematic Analysis of Knowledge* (pp. 189–211). Springer. https://doi.org/10.1007/978-1-4419-5662-0_11

Scriven, M. (1980). Prescriptive and descriptive approaches to problem solving. In D. T Tuma and F. Reif, (Eds.), *Problem solving and education: Issues in teaching and research* (127-139). Lawrence Erlbaum Associates.

Senge, P., Kleiner, A., Roberts, C., Ross, R., & Smith, B. (1994). *The fifth discipline fieldbook: Strategies and tools for building a learning organization*. Doubleday.

Sheppard, L. D., & Vernon, P. A. (2008). Intelligence and speed of information processing: A review of 50 years of research. *Personality and Individual Differences, 44*(3), 535-551. https://doi.org/10.1016/j.paid.2007.09.015

Sill, D. J. (1996). Integrative thinking, synthesis, and creativity in interdisciplinary studies. *The Journal of General Education*, *45*(2), 129-151. http://www.jstor.org/stable/27797296

Simon, D. P., & Simon, H. A. (1978). Individual differences in solving physics problems. In R. S. Siegler (Ed.), *Children's thinking: What develops?* (pp. 325–348). Lawrence Erlbaum Associates.

Simon, H., & Kotovsky, K. (1963). Human acquisition of concepts for sequential patterns. *Psychological Review, 70*(6), 534–546. https://doi.org/10.1037/h0043901

Skaržauskienė, A. (2008). Theoretical insights to leadership based on systems thinking principles. *Organizacijų Vadyba: Sisteminiai Tyrimai*, *48*, 105-120. https://www.ceeol.com/search/article-detail?id=135652

Skaržauskienė, A. (2009). Systems thinking as a competence in the leadership paradigm. *Management Theory and Studies for Rural Business and Infrastructure Development*, *16*(1), 97-105.

Skaržauskienė, A. (2010). Managing complexity: systems thinking as a catalyst of the organization performance. *Measuring Business Excellence, 14*(4), 49-64. https://doi.org/10.1108/13683041011093758

Skaržauskienė, A., & Jonušauskas, S. (2013). Defining innovative competencies of leading: systems intelligence. *International Journal of Innovation and Regional Development 5*(2), 132-148. https://doi.org/10.1504/IJIRD.2013.055236

Sojka, J. Z., & Giese, J. L. (2006). Communicating through pictures and words: Understanding the role of affect and cognition in processing visual and verbal information. *Psychology & Marketing, 23*(12), 995-1014. https://doi.org/10.1002/mar.20143

Sok, P., & O'Cass, A. (2015). Examining the new product innovation–performance relationship: Optimizing the role of individual-level creativity and attention-to-detail. *Industrial Marketing Management*, *47*, 156-165. https://doi.org/10.1016/j.indmarman.2015.02.040

Sterman, J. D. (1989). Modeling managerial behavior: Misperceptions of feedback in a dynamic decision-making experiment. *Management Science. 35*(3), 321–339. https://doi.org/10.1287/mnsc.35.3.321

Sternberg, R. J. (1986). *Intelligence applied: Understanding and increasing your intellectual skills.* Harcourt Brace Jovanovich.

Sternberg, R. J. (2009). *Cognitive psychology*. Wadsworth/Cengage Learning.

Stokols, D. (2006). Toward a science of transdisciplinary action research. *American Journal of Community Psychology*, *38*(1), 63-77. https://doi.org/10.1007/s10464-006-9060-5

Stowe, D. E., Eder, D. J., Wentworth, J., & Sebberson, D. (2002). Interdisciplinary program assessment. *Issues in Interdisciplinary Studies*, *20*, 77-101. http://hdl.handle.net/10323/4392

Streufert, S., & Nogami, G. Y. (1989). Cognitive style and complexity: Implications for I/O psychology. In C. L. Cooper & I. T. Robertson (Eds.), *International review of industrial and organizational psychology* (pp. 93–143). John Wiley & Sons.

Suedfeld, P., & Tetlock, P. E. (1977). Integrative complexity of communications in international crises. *Journal of Conflict Resolution, 21*(1), 169-184. https://doi.org/10.1177/002200277702100108

Suedfeld, P., Tetlock, P. E., & Streufert, S. (1992). Conceptual/integrative complexity. In C. P. Smith, J. W. Atkinson, D. C. McClelland, & J. Veroff (Eds.), *Motivation and personality: Handbook of thematic content analysis* (pp. 393-400). Cambridge University Press.

Tippins, M. J., & Sohi, R. S. (2003). IT competency and firm performance: Is organizational learning a missing link?. *Strategic Management Journal, 24*(8), 745-761. https://doi.org/10.1002/smj.337

Ungson, G. R., Braunstein, D. N., & Hall, P. D. (1981). Managerial information processing: A research review. *Administrative Science Quarterly 26*(1), 116-134. https://doi.org/10.2307/2392604

U.S. Department of the Army (2019). *Army Leadership and the Profession* (Army Doctrine Publication 6-22). Author.

Van Camp, L., Sabbe, B., & Oldenburg, J. (2017). Cognitive insight: A systematic review. *Clinical Psychology Review, 55,* 12-24. https://doi.org/10.1016/j.cpr.2017.04.011

Van Deursen, A., Van Dijk, J., & Peters, O. (2012). Proposing a Survey Instrument for Measuring Operational, Formal, Information and Strategic Internet Skills. *International Journal of Human-Computer Interaction, 28*(12), 827-837. https://doi.org/10.1080/10447318.2012.670086

Van Deursen, A.J.A.M., & Van Dijk, J.A.G.M. (2009a). Using the Internet: Skill Related Problems in Users' Online Behavior. *Interacting with Computers, 21*(5/6), 393-402. https://doi.org/10.1016/j.intcom.2009.06.005

Van Deursen, A.J.A.M., & Van Dijk, J.A.G.M. (2009b). Improving digital skills for the use of online public information and services. Government Information Quarterly, *26*(2), 333-340. https://doi.org/10.1016/j.giq.2008.11.002

Van Deursen, A.J.A.M., & Van Dijk, J.A.G.M. (2010). Measuring Internet skills. *International Journal of Human-Computer Interaction, 26*(10), 891-916. https://doi.org/10.1080/10447318.2010.496338

Van Deursen, A.J.A.M., & Van Dijk, J.A.G.M. (2011). Internet Skills and the Digital Divide. *New Media & Society, 13*(6), 893-911. https://doi.org/10.1177/1461444810386774

Van Deursen, A., Helsper, E .J. & Eynon, R. (2014). *Measuring Digital Skills. From Digital Skills to Tangible Outcomes* (Project Report). Oxford Internet Institute. http://www.oii.ox.ac.uk/research/projects/?id=112

Van Deursen, A.J.A.M., Helsper, E. J., & Eynon, R. (2016). Development and validation of the Internet Skills Scale. *Information, Communication & Society, 19*(6), 804-823, https://doi.org/10.1080/1369118X.2015.1078834.

Van Dijk, J.A.G.M., & Van Deursen, A.J.A.M. (2014). *Digital skills, unlocking the information society.* Palgrave Macmillan.

Venkatesh, V., & Davis, F. D. (2000). A theoretical extension of the technology acceptance model. *Management Science, 46*(2), 186– 204. https://doi.org/10.1287/mnsc.46.2.186.11926

Vernon, D., Hocking, I., & Tyler, T. C. (2016). An evidence-based review of creative problem solving tools: A practitioner's resource. *Human Resource Development Review*, *15*(2), 230-259. https://doi.org/10.1177/1534484316641512

Vester, F. (2010). *Ecopolicy: A Cybernetic Environmental Simulation Game*. English/German version. Malik.

Vincenti, V. B. (2001). Exploration of the relationship between international experiences and the interdisciplinary work of university faculty. *Journal of Studies in International Education*, *5*(1), 42-63. https://doi.org/10.1177/102831530151004

Walsh, J. P. (1995). Managerial and organizational cognitions: Notes from a trip down memory lane. *Organization Science 6*(3), 280-321. https://doi.org/10.1287/orsc.6.3.280

Wang, R. Y., & Strong, D. M. (1996). Beyond accuracy: What data quality means to data consumers. *Journal of Management Information Systems, 12*(4), 5-33. https://doi.org/10.1080/07421222.1996.11518099

Wang, Z., & Tchernev, J. M. (2012). The "myth" of media multitasking: Reciprocal dynamics of media multitasking, personal needs, and gratifications. *Journal of Communication*, *62*(3), 493-513. https://doi.org/10.1111/j.1460-2466.2012.01641.x

Watson, J. M., & Strayer, D. L. (2010). Supertaskers: Profiles in extraordinary multitasking ability. *Psychonomic Bulletin & Review*, *17*(4), 479-485. https://doi.org/10.3758/PBR.17.4.479

Watson, M. (2018, August 10). *5 reasons we created the Opex Analytics beer game.* Opex-Analytics. https://medium.com/opex-analytics/5-reasons-we-created-the-opex-analytics-beer-game-852b434b9aa7

Wetherell, M. A., & Carter, K. (2014). The multitasking framework: The effects of increasing workload on acute psychobiological stress reactivity. *Stress and Health*, *30*(2), 103-109. https://doi.org/10.1002/smi.2496

Wilhelm, O. (2005). Measuring reasoning ability. In R Engle, O Wilhelm (Eds.), *Handbook of Understanding and Measuring Intelligence* (pp. 373-392). Sage Publications.

Wirth, J., & Klieme, E. (2003). Computer-based assessment of problem solving competence. *Assessment in Education: Principles, Policy & Practice*, *10*(3), 329-345. https://doi.org/10.1080/0969594032000148172

Zaretsky, L. (2007). A transdisciplinary team approach to achieving moral agency across regular and special education in K-12 schools. *Journal of Educational Administration*, *45*(4), 496-513. https://doi.org/10.1108/09578230710762472

Żelechowska, D., Żyluk, N., & Urbański, M. (2020). Find out a new method to study abductive reasoning in empirical research. *International Journal of Qualitative Methods*, *19*, 1-11. https://doi.org/10.1177/1609406920909674

Appendix A

Analytical Thinking

<table>
<tr><th colspan="2">Measurement Details</th></tr>
<tr><td>Construct: Analytical Thinking</td><td>Subdimensions: Information Evaluation, Hypothesis Evaluation, Process Evaluation</td></tr>
<tr><td colspan="2">Source (ATAF) Definitions:

Analyze Data or Information: Identifies underlying principles, relationships, general rules, reasons, or facts by breaking down information or data into separate parts

Analytical Thinking: Analyzes information and applies general rules and logic to address work-related issues and problems</td></tr>
<tr><td colspan="2">Operational Definition:

Analyzes information and applies general rules and logic to evaluate the information at hand, which is often incomplete, and comes to the best possible conclusion or decision.</td></tr>
<tr><td colspan="2">Decisions:

• Focus only on information, not data (qualitative analysis over quantitative)
• Focus on abductive reasoning
• Do not include hypothesis generation (a part of the abductive reasoning process) in the measure as would require a constructed response question; for ease of programming and scoring, this will not be measured in the current assessment</td></tr>
<tr><td colspan="2">Measurement Overview: Test takers complete a single measure that examines two analytic constructs (Analyze Data or Information and Analytical Thinking). In order to complete this measure, test takers must evaluate information and engage in abductive reasoning (a person’s ability to develop and evaluate logical inferences or hypotheses in the absence of complete information) while completing four different item types.</td></tr>
<tr><td colspan="2">Measure Details:

Description:

• Item Type 1: Item stems include an incomplete set of observations or a short scenario. Test takers are asked the extent to which the information provided is sufficient to determine an explanation for what happened in the scenario (1=not at all sufficient to 5=completely sufficient). Items will be designed to vary with regard to whether there is sufficient information to determine an explanation. This item type is designed to measure the information evaluation subdimension by assessing how well test takers evaluate the sufficiency of the information at hand.

• Item Type 2: Regardless of whether there is sufficient information, test takers are instructed to select the most plausible explanation for the scenario from response options that list possible explanations or hypotheses. For items with sufficient information, the correct response is given the most points. For items will insufficient information, the most likely explanation is given the</td></tr>
</table>

most points. There may be other explanations, but test takers are being asked to choose the most plausible among the possibilities presented. This item type is designed to measure the hypothesis evaluation subdimension by assessing how well test takers evaluate the plausibility of different hypotheses.

- Item Type 2 (follow up): Item stems provide additional information about a scenario and then ask test takers to select the most plausible response of those given. This item type is designed to measure the hypothesis evaluation subdimension by assessing how well test takers evaluate the plausibility of different hypotheses. Item 3a below is an example of an item where a test taker's hypothesis should not change based on the information provided, while item 3b is an example where a test taker's hypothesis should change based on the information provided.
- Item Type 3: Item stems ask test takers which question they would pose first to determine what has happened within the given scenario. This item type is designed to measure the process evaluation subdimension by assessing how well test takers evaluate the reasoning process for gathering additional information to refine hypotheses.
- Item Type 4: Item stems provide test takers with a statement that takes a stance on a certain issue and then ask them to identify the response options that represent the "most" and "least" effective counterpoints. This item type is designed to measure the information evaluation subdimension by assessing how well test takers evaluate the sufficiency of the information at hand.

Scoring:

Empirical Keying

Item developers will have the "best" answer in mind when writing the items. However, due to the potential lack of one, and only one, correct answer (much like an SJT item), an empirical scoring key would be considered. This would involve the following steps:

1. Split criterion-validation sample into a validation sample (80% of all test takers) and a holdout sample (20% of all test takers).
2. Compute average test taker scores on criterion of interest (e.g., Officer job/training performance).
3. Score response options such that response options associated with the highest criterion score (i.e., test takers who endorsed this response option had the highest criterion score when compared to those who endorsed other response options) are given a value of 1. The response option with the lowest criterion score would be given a value of 0, and any response(s) associated with a criterion score between the highest and lowest would be given a value of .67, .5 or .33 (depending on the number of responses that fell between the highest and lowest scores).
4. Compute biserial correlations between the rescored items and the criterion scores for both the validation and hold out samples.
5. Items that meet a certain threshold in both samples will be retained.

Rational Scoring

An alternative approach would be to use a rational based scoring key in which response options are assigned a value based on input by the test developers and SMEs. In this case, the response option determined to be the most plausible option would receive point, the least plausible option which would receive zero points, and the other response options would be assigned partial credit according to how relatively plausible they are.

This scoring strategy could be administered in CAT framework to reduce item exposure and testing time while increasing the precision of the ability estimate. However, in this case, it would not be possible to

use questions that rely on other information contained in other questions for the test taker to be able to answer. All item types could be used if each item included an independent scenario.

Example Items:

You arrive at the airport and make it through security. Before heading to your gate, you make your way through the crowded food court towards a restroom. As you approach the restroom, you see a black bag sitting on a bench. No one else is sitting on the bench where the bag is, and you look at the people walking around the area but don't see anyone who appears to own the bag.

1. Rate the extent to which the information provided is sufficient to determine an explanation for what happened in the scenario above.

- Likert scale: 1=not at all sufficient to 5=completely sufficient

2. Select the most plausible explanation for why the bag is sitting on the bench.

A. Someone inadvertently forgot their bag
B. Someone left their bag on the bench while they got food
C. It is there as part of a terrorist attack
D. It was stolen then discarded there
E. Someone is reserving a seat with it

3a. As you enter the restroom, you hear a security announcement warning all passengers to keep an eye on their bags and report any suspicious activity immediately to security guards. Select the most plausible explanation for why the bag is sitting on the bench.

A. Someone inadvertently forgot their bag
B. Someone left their bag on the bench while they got food
C. It is there as part of a terrorist attack
D. It was stolen then discarded there
E. Someone is reserving a seat with it

3b. As you are leaving the restroom, a parent pushing a stroller with two kids running around are in front of you. The parent is attempting to herd the kids back through the area, while pushing the stroller and carrying a diaper bag. She is looking around frantically at the area outside of the restroom as if she is searching for something or someone. Select the most plausible explanation for why the bag is sitting on the bench.

A. Someone inadvertently forgot their bag
B. Someone left their bag on the bench while they got food
C. It is there as part of a terrorist attack
D. It was stolen then discarded there
E. Someone is reserving a seat with it

4. What question would you ask first to determine what has happened?

A. How long has the bag been unattended?
B. Who is the owner of the bag?
C. What are the contents of the bag?
D. Where is the owner of the bag?
E. What is the owner of the bag currently doing?

Please read the following statement that provides an opinion and justification about a given topic. Then consider the effectiveness of the each of the counterpoints that are offered and select the one that is MOST effective and LEAST effective as a counterpoint to the justification offered.

5. College athletes do not need to be compensated beyond receiving a scholarship because they are also receiving a free education, which is priceless. Indicate the MOST and LEAST effective counterpoints.

A. Even with a full scholarship (covering tuition, fees, room, board, and books), an athlete will have to pay money out of pocket to bridge the cost-of-living gap.
B. Most sports scholarships are not guaranteed. They are renewable yearly and can be terminated at any time, including for injuries.
C. Student athletes tend to be over-scheduled with study halls, practices, weight training sessions, film study, individual workouts, travel, and competition in an attempt to help athletes maintain focus on their sport.
D. Some college athletes make their schools millions of dollars. It is only fair that they receive a cut of the profits.
E. Participating in intercollegiate athletics constitutes a full-time job and athletes should be compensated accordingly.

Appendix B

Detailed Information Processing

Measurement Details	
Construct: Detailed Information Processing	**Subdimensions:** None

Source (ATAF) Definition:

Processes Information or Data: Compiles, codes, categorizes, calculates, tabulates, audits, or verifies information or data

Component Definitions (merriam-webster.com/dictionary, 2021)

- Compile: to compose out of materials from other documents
- Code: to put in or into the form or symbols of a code (code = a system of signals or symbols for communication)
- Categorize: to put into a category, classify (category = a division within a system of classification)
- Tabulate: to count, record, or list systematically
- Audit: a methodical examination and review
- Verify: to establish the truth, accuracy, or reality of the claim

Detail Focused: Attentive to detail and thorough, accurate, and precise in completing a task

Operational Definition:

Attentive to detail and thorough, accurate, and precise in processing (compiling, categorizing, auditing, and verifying) information

Decisions:

- Focus on specific skills listed in subdimensions of processes information or data (e.g., audits information)
- Focus on cognitive aspects of constructs rather than trait-based aspects
- Combine the original Processes Information and Data and Detail-Focused constructs due to considerable overlap in the conceptual constructs and the operational measurement options
- Include the different skills listed under processes information as "components" rather than "subdimensions" because we are not trying to produce scores for each component (thus, distinguishing them from subdimensions)

Measurement Overview: A single measure for the two constructs that uses military-type objects like reports or simple graphics, asking test takers to identify errors, problems, etc.

Measure Details:

Description:

Task-based assessment where multiple data points are collected, increasing the number of measurement opportunities to allow for a range of difficulty within and across items.

Scoring:

This format could be scored using a traditional scoring method (points for correct answers and no points for incorrect answers) or a Generalized Partial Credit Model (GPCM), which allows for partial credit scoring within each scenario. GPCM generates more item information and leads to lower standard error at the test level, increasing test reliability.

This approach also allows for a CAT administration. However, it would require generating multiple scenarios/items to administer within this framework. A traditional scoring approach could be used within one scenario, which would require less administration time, but could not be used in a CAT framework because the items are not independent (they all rely on the same scenario).

This can be administered as multiple items within one scenario with a traditional scoring approach, or as multiple item parts across multiple scenarios with a GPCM scoring approach in a CAT framework to reduce item exposure, while increasing the precision of the ability estimate.

Example Items:

Note: Correct response options are bolded.

Recommendation for Award Form (DA Form 638)
Your supervisor has asked you to complete a DA 638 form to recommend a Soldier for an award. A DA 638 form is used by the U.S. Army and is completed when a higher-ranking Soldier wishes to recommend another Soldier for an award. This form requires information about the Soldier that is being considered for the award as well as information about the office that is making the recommendation. Your supervisor provided the following information and has asked you to fill in the details on the form.

> Over the span of two years, while serving at Fort Gordon, from 02 January 2019 to 28 February 2021, Sergeant Bolton managed and utilized a $500,000 video teleconference system to present a daily intelligence brief to the agency's Director and other key leaders. He provided a total of 1,400 conferences and 500 hours of global communications support to over 60 internal and external agency sites. In addition, Sergeant Bolton led the military desktop support team in the creation of 177 network accounts, the resolution of 334 trouble tickets and the restoration of 22 high priority issues. His diligence and dedication to duty reflects great credit upon Cyber Operations and the United States Army.

Compile & Detail Focused:

1. Fill in the available fields based on the information your supervisor and the instructions provided.

Categorize:

2. Which type(s) of award(s) do you think Sergeant Bolton should receive, based on the work your supervisor described and the descriptions of the awards? [Select all that apply]

- Medal of Honor
- Distinguished Service Cross
- **Distinguished Service Medal**
- Silver Star
- **Legion of Merit**
- Distinguished Flying Cross
- Soldier's Medal
- Bronze Star Medal

- **Meritorious Service Medal**
- Air Medal

Audit:

3. What additional information do you need to request to complete the form? [Select all that apply]
- RANK
- **DODID/SSN**
- **ORGANIZATION**
- **PREVIOUS AWARDS**
- PERIOD OF AWARD

Audit & Verify:

4. A coworker asks you to review a form they completed to make sure the content is correct. Which fields were completed incorrectly? [Select all that apply]
- 4. NAME
- 9. BRANCH OF SERVICE
- 10. RECOMMENDED AWARD
- 11. PERIOD OF AWARD
- 20. ACHIEVEMENTS

RECOMMENDATION FOR AWARD

For use of this form, see AR 600-8-22; the proponent agency is DCS, G-1.

For valor/heroism/wartime and all awards higher than MSM, refer to special instructions in Chapter 3, AR 600-8-22.

1. TO	2. FROM	3. DATE (YYYYMMDD)
Commander, 82d Airborne Division, Ft Bragg, NC 28310	Commander, HHD, 27th Main Support Bn, 1st Cavalry Division, Ft. Hood, TX 76544	20060710

PART I - SOLDIER DATA

4. NAME (Last, First, Middle Initial)	5. RANK	6. SSN
DOE, JOHN B Sr.	SSG	000-11-2222

7. ORGANIZATION	8. PREVIOUS AWARDS		
HHD, 27th Main Support Bn, 1st Cavalry Division, Ft. Hood, TX 76544	AAM-4, ARCOM-2, MSM-1, BSM-1		
9. BRANCH OF SERVICE	10. RECOMMENDED AWARD	11. PERIOD OF AWARD	
		a. FROM	b. TO
	MSM 1OLC	20020401	20061009

12. REASON FOR AWARD

12a. INDICATE REASON	12b. INTERIM AWARD YES [ ] NO [X] IF YES, STATE AWARD GIVEN	12c. POSTHUMOUS YES [ ] NO [X]	13. PROPOSED PRESENTATION DATE (YYYYMMDD)
PCS			20061212

PART II - RECOMMENDER DATA

14. NAME (Last, First, Middle Initial)	15. ADDRESS
JONES, YOLANDA A.	HHD, 27th Main Support Bn Ft. Hood, TX 76544
16. TITLE/POSITION: Platoon Sergeant — 17. RANK: SFC	
18. RELATIONSHIP TO AWARDEE: Platoon Sergeant	19. SIGNATURE: //SIGNED//

PART III - JUSTIFICATION AND CITATION DATA *(Use specific bullet examples of meritorious acts or service)*

20. ACHIEVEMENTS

ACHIEVEMENT #1

SSG Doe was responsible for all logistics matters pertaining to the unit. He reviewed internal logistical status reports and provided daily logistical information to the higher headquarters.

ACHIEVEMENT #2

SSG Doe maintained the current status of the commander's critical list. He was responsible for coordinating transportation requests for the administrative moves.

ACHIEVEMENT #3

SSG Doe was responsible for ensuring highway clearances were submitted and supervised over internal supply and maintenance procedures. He provided support to all battalion-level command, staff personnel, and battalion food service personnel.

ACHIEVEMENT #4

SSG Doe provided support to the units Forward Support Battalions and the Aviation Support Battalion.

21. PROPOSED CITATION

For exceptionally meritorious service while assigned as Noncommissioned Officer in Charge of the Support Operations Section. Staff Sergeant Doe's leadership and selfless devotion to duty enhanced operational support and ensured mission accomplishment. His professionalizm reflects great credit upon himself, the 1st Cavalry Division, and the United States Army.

DA FORM 638, APR 2006 — REPLACES DA FORM 638-1. PREVIOUS EDITIONS OF DA FORM 638 ARE OBSOLETE. — Page 1 of 3 — APD PE v2.01ES

Provided materials:

Instructions

Block 1. Fill in the office information of the next office above the recommender in the Chain of Command. This is the office with approval authority.

Block 2. Enter the recommender's office information.

Blocks 3-7. Self-Explanatory

Block 8. Fill in the individual's previous achievement awards. The Army Service Ribbon (ASR), Korean Defense Service Medal (KDSM), Kuwait Liberation Medal (KULIBM), Southwest Asia Service Medal (SWASM) are service medals and should not be listed.

Block 9. For Soldiers in the U.S. Army, leave this block blank. Make an entry only if recommending someone from another service or a foreign military member for an award. For members of the other U.S. Armed Forces, enter the Service (for example, U.S. Air Force). For foreign military personnel enter the country (for example, Federal Republic of Germany). See AR 600-8-22 Table 3-2.

Block 10. Enter the award being recommended to include number of award/oak leaf cluster

Block 11-19. Self-Explanatory

Block 20. For the AAM, ARCOM, or MSM, using the space provided in the DA 638, fill in the Soldier's achievements or meritorious service using bullet format. For the BSM and above, type up a separate narrative (see below).

Narratives for valor must contain a description of the following elements:

A. Terrain and weather of the area in which the action took place.
B. Enemy conditions, to include morale, proximity, firepower, casualties, and the circumstances before, during, and after the act.
C. The effect of the act on the enemy.
D. The action of comrades in the immediate vicinity of the act and the degree of their participation.
E. If the act occurred in flight, the type and position of the aircraft and the individual's crew position
F. The degree to which the act was voluntary.
G. The degree to which the act was outstanding and exceeded what was normally expected of the individual.
H. All unusual circumstances and overall effects or results of the act.

Recommendations for awards for heroism or valor must contain witness statements, preferably in the form of affidavits or sworn statements. If possible, recommendations should be supported by official records, maps, photographs, etc.

Block 21. For an AAM, ARCOM, or MSM, enter the citation. For a BSM/AM or higher, write the appropriate citation (see below).

Block 22. Staff Section Admin Clerk should sign and date verifying individual is eligible for the recommended award.

Block 23-25. Fill in intermediate commanders' as appropriate.

Block 26.

A. ORDERS ISSUING AUTHORITY
B. CDR, MNC-I APO AE 09999
C. JOHN Q. PUBLIC
D. COL
E. COMMANDING

Block 27.

A. HQ, MNC-I APO AE 09999

Block 28.

A. JOSEPH M. BLOW

B. MSG
C. Chief, CJ1 Personnel Actions

Block 29-30. Leave blank

Block 31.
1-SOLDIER
1-UNIT
1-FILE
1-OMPF

Award Descriptions

Medal of Honor: awarded to a person who, while a member of the Army, distinguishes himself or herself conspicuously by gallantry and intrepidity at the risk of his or her life above and beyond the call of duty while engaged in an action against an enemy of the United States; while engaged in military operations involving conflict with an opposing foreign force; or while serving with friendly foreign forces engaged in an armed conflict against an opposing armed force in which the United States is not a belligerent party. The deed performed must have been one of personal bravery or self-sacrifice so conspicuous as to clearly distinguish the individual above his or her comrades and must have involved risk of life. Incontestable proof of the performance of the service will be exacted and each recommendation for the award of this decoration will be considered on the standard of extraordinary merit.

Distinguished Service Cross: awarded to a person who, while serving in any capacity with the Army, distinguished himself or herself by extraordinary heroism not justifying the award of a MOH; while engaged in an action against an enemy of the United States; while engaged in military operations involving conflict with an opposing or foreign force; or while serving with friendly foreign forces engaged in an armed conflict against an opposing armed force in which the United States is not a belligerent party. The act or acts of heroism must have been so notable and have involved risk of life so extraordinary as to set the individual apart from their comrades.

Distinguished Service Medal: awarded to any person who, while serving in any capacity with the U.S. Army, has distinguished himself or herself by exceptionally meritorious service to the government in a duty of great responsibility. The performance must be such as to merit recognition for service that is clearly exceptional. Exceptional performance of normal duty will not alone justify an award of this decoration.

Silver Star: awarded to a person who, while serving in any capacity with the U.S. Army, is cited for gallantry in action against an enemy of the United States while engaged in military operations involving conflict with an opposing foreign force, or while serving with friendly foreign forces engaged in armed conflict against an opposing armed force in which the United States is not a belligerent party. The required gallantry, while of a lesser degree than that required for the DSC, must nevertheless have been performed with marked distinction.

Legion of Merit: awarded to any Servicemember of the Armed Forces of the United States or a friendly foreign nation who has distinguished himself or herself by exceptionally meritorious conduct in the performance of outstanding services and achievements.

Distinguished Flying Cross: awarded to any person who, while serving in any capacity with the Army of the United States, distinguished himself or herself by valor, heroism, or extraordinary achievement while participating in aerial flight.

Soldier's Medal: awarded to any person of the Armed Forces of the United States or of a friendly foreign nation who, while serving in any capacity with the Army of the United States, including Reserve Component Soldiers not serving in a duty status, as defined in 10 USC 101(d), at the time of the heroic act, who distinguished himself or herself by heroism not involving actual conflict with an enemy.

Bronze Star Medal: awarded to any person who, while serving in any capacity in or with the Armed Forces of the United States, or a friendly foreign nation, after 6 December 1941, distinguished himself or herself by heroic or meritorious achievement or service, not involving participation in aerial flight, in connection with military operations against an armed enemy; or while engaged in military operations involving conflict with an opposing armed force in which the United States is not a belligerent party. 10 USC 1133 limits award of the BSM to Servicemembers receiving imminent danger pay, hostile fire pay or hazardous duty pay and members of a friendly military force who are serving in a geographic area in which special pay is authorized under 37 USC 310 or 37 USC 351(a) (1) and (3) or special pay under any of the following circumstances:

(1) While engaged in an action against an enemy of the United States.

(2) While engaged in military operations involving conflict with an opposing foreign force.

(3) While serving with friendly foreign forces engaged in an armed conflict against an opposing armed force in which the United States is not a belligerent party.

Meritorious Service Medal: awarded to any Servicemember of the Armed Forces of the United States or to any member of the armed forces of a friendly foreign nation who has distinguished himself or herself by outstanding meritorious achievement or service.

Air Medal: awarded to any person who, while serving in any capacity in or with the Armed Forces of the United States, who has distinguished himself or herself by meritorious achievement while participating in aerial flight. Awards may be made to recognize single acts of valor, merit or heroism, or for meritorious service as described below:

(1) Awards may be made for acts of heroism in connection with military operations against an armed enemy or while engaged in military operations involving conflict with an opposing armed force in which the United States is not a belligerent party, which are of a lesser degree than required for award of the DFC.

(2) Awards may be made for a single act of meritorious achievement, involving superior airmanship, which are of a lesser degree than required for award of the DFC, but nevertheless were accomplished with distinction beyond that normally expected.

(3) Awards for meritorious service may be made for sustained distinction in the performance of duties involving regular and frequent participation in aerial flight for a period of at least 6 consecutive months (a month is considered 30 calendar days). In this regard, accumulation of a specified number of hours and missions will not serve as the basis for award of the AM. Criteria in paragraph 3–18b(1) concerning conditions of conflict are applicable to award of the AM for meritorious service.

(4) Awards for meritorious, deployed service of six cumulative months, with no limitations on the number of deployments within a 24-month period for certain Special Operations Aviation and Military Intelligence Aerial Reconnaissance units with an exception to policy approved by The Adjutant General. Exception to policy requests will be submitted to HRC, ADB with full justification and supporting documentation.

DA 638 Form: See separate attachment

Additional Materials (not currently used to address items)

Awards Table:
This table could be used to generate another item to tap "categorizes". E.g., Which of the following awards are awarded only for service during combat conditions?

- Medal of Honor
- Distinguished Service Cross
- Distinguished Service Medal
- Silver Star
- Legion of Merit
- Distinguished Flying Cross
- Soldier's Medal
- Bronze Star Medal
- Meritorious Service Medal
- Air Medal
- Army Commendation
- Army Achievement Medal

Individual Awards and Authorized Uses

Title of Decoration	Valor only	Valor "V" device	Achievement under combat conditions. "C" device	Remote combat impacts "R" device	Noncombat heroism	Other specific achievement	Other meritorious service
Medal of Honor[1]	X						
Distinguished Service Cross[1]	X						
Distinguished Service Medal			X			X	X
Silver Star[1]	X						
Legion of Merit			X	X		X	X
Distinguished Flying Cross		X	X		X	X	
Soldier's Medal[2]					X		
Bronze Star Medal[3]		X					
Purple Heart[4]	N/A	N/A	N/A	N/A	N/A	N/A	N/A
Meritorious Service Medal[5]				X		X	X
Air Medal		X	X		X	X	X
Army Commendation Medal		X	X	X	X	X	X
Army Achievement Medal			X	X		X	X

Notes:

[1] These awards are by definition valor awards. The "V" and "C" devices are not authorized.

[2] The "V" and "C" devices are not authorized; The Soldier's medal is only awarded for non-combat heroism.

[3] The "C" device is not authorized because the BSM is a combat related award and service or achievement under combat conditions is inherent to the medal. All awards for non-valorous achievement or service must satisfy the requirements for personal exposure to hostile action or significant risk of exposure to hostile action.

[4] For wounds received as a result of hostile action. No combat devices are authorized.

[5] The "V" and "C" devices are not authorized. The MSM may be awarded for outstanding achievement or service in a combat zone, combat zone tax exclusion area and areas authorized for imminent danger pay, hostile fire pay, or hazardous duty pay if the achievement or service is not combat related. The BSM is the appropriate equivalent level award to recognize meritorious achievement or service under combat conditions.

The "V" device denotes participation in acts of heroism involving conflict with an armed enemy. The “C” device is only authorized if the service or achievement was performed while the Servicemember was personally exposed to hostile action or was at significant risk of exposure to hostile action. The “R” device denotes the direct hands-on employment of a weapon system or other warfighting activities that had a direct and immediate impact on a combat operation or other military operation (that is, outcome of an engagement or specific effects on a target).

Appendix C

Multitasking

Measurement Details	
Construct: Multitasking	**Subdimensions:** None

Source (ATAF) Definition:

Rapidly processes and prioritizes multiple demands simultaneously; Takes appropriate action when multiple stimuli compete for his or her attention

Operational Definition*:*

The ability to shift attention between two or more tasks with minimal reduction in performance

Decisions:

- Remove "takes appropriate action" from the definition
- Focus on the attention-shifting aspect of the construct for this measure
- ARI preferences for the assessment are for low complexity, low computer resources, and high test security (so that multiple retakes are possible and do not show practice effects)
- The term "simultaneously" is being interpreted as within the same timeframe, not literally at the same time

Measurement Overview: Test takers complete a visual search task to establish a baseline score, then complete the same task while interruptive prompts are occurring to measure the level of performance decrement due to shifting tasks.

Measure Details:

Description:

Test takers will first be provided with a visual search task where they are asked to select Soldiers with certain characteristics or qualifications from a list so that these Soldiers can be assigned to a special unit. Test takers are provided with a table that lists Soldiers and descriptions of certain characteristics such as MOS, rank, and time-in-grade, as well as a set of rules regarding how the selections should be made. In order to prevent an increase in performance due to memorization of Soldier qualifications, for each new item, a new list of Soldiers with different qualifications will be provided or the names in the list will be shuffled and names already used will be replaced. Test takers will be asked to select the appropriate Soldiers by double clicking the names that meet the stated requirements. Test takers are given 10 seconds per item to find the appropriate Soldiers before the test advances to the next item. To increase variance, each list will contain multiple names that meet the requirements, requiring test takers to identify as many appropriately qualified Soldiers as they can within the timeframe.

An initial set of 8-12 items will be completed without interruption tasks and will serve as the baseline for performance. The number of items in the initial set will be determined based on the amount of time each item takes to process, with a goal of achieving a baseline time-period of around two minutes. After establishing a baseline score, test takers will complete another set of approximately 28 items, but this

time will also receive interruption tasks every 3-5 items. These tasks will be in the form of pop-up windows made to look like messages from a colleague or superior asking for information about a Solider (e.g., whether they are deployable or what their specific qualifications are) or general information about the set of Soldiers (e.g., are there at least five deployable Infantry Soldiers in your current list?). For these tasks it will be necessary to ensure the test taker is able to look at the list of Soldiers so that they do not have to rely on working memory to answer the questions.

Test takers will provide a Yes/No response to the interruption task, then return to the visual search task. While the interruption task is on screen, only the table from the visual search task will be visible – not the Soldier specification directions used to make the selections for the special unit, and the test taker will not be able to engage in selecting responses for the visual search task. The interruption task will remain on screen until it has been answered or 10 seconds has passed (time to be determined), whichever comes first. At that time, the next visual search task item will appear. For the trial to count, the test taker must answer the interruption task correctly. Additional items are added for each incorrect response. The test ends four items after the participant answers five interruption tasks correctly.

Scoring:

Formula or Sensitivity Metric

Scores are based on search performance, using either a formula scoring approach that adds points for each correct click (hits) and subtracts points for each incorrect click (false alarms) or a sensitivity metric like d-prime which accounts for hits, misses, false alarms, and correct rejections. If a sensitivity metric is used, several correct selections will be required per table. Two scores -a baseline and a multitasking score- are calculated as follows:

1. Baseline Score - An initial set of 8-12 items will be completed without interruptions and will serve as the baseline for performance. Responses will be scored based on the number correct. Test takers can achieve partial credit by selecting only some of the correct Soldiers for the special unit, which should increase the range of scores and help differentiate between high and low performers more effectively. An overall score for the initial items is calculated as the baseline score.
2. Multitasking Score – A second set of items will be completed with periodic interruptions that must be addressed before the test taker can return to the original task. Responses will be scored using the same procedure as the baseline. An overall score for the interruption trials is calculated and the performance decrement between the baseline and multitasking conditions is used to calculate the final Multitasking score. A smaller performance decrement between the multitasking and the baseline conditions would result in a higher final multitasking score.

Discontinuous Growth Modeling

An alternative scoring approach that could be considered would be to use discontinuous growth modeling to examine performance trajectories for each test taker. This approach allows for the calculation of more nuanced performance metrics, such as rate of recovery several trials after an interruption, as well as reactions to the interruption, or even correlation between performance before the interruption and the degree of decrement caused by interruption.

Example Items:

For this task, you are being asked to assign Soldiers to special units based on their qualifications. For each question you will be provided with the specific requirements for that special unit and a table of Soldiers from which to choose. You must examine the requirements for the special unit and then select the Soldiers that meet the requirements by double-clicking on their name in the table.

As you complete this task, you may receive messages from various coworkers asking for specific information about the Soldiers in the table. Please respond to these messages immediately and to the best of your ability, then continue with the special unit assignment task.

This task will take between 5-10 minutes.

1. Please select all Soldiers in Roster 1 who have a rank of 1LT and MOS of 11A. [Correct responses can be seen in bold.]

Roster 1				
Soldier	**MOS**	**Rank**	**Time in Grade**	**Deployable**
1. Gonzalez	**11A**	**1LT**	**6 months**	**No**
2. Kim	12A	LTC	12 months	Yes
3. Graham	25A	CPT	18 months	No
4. Mendez	25A	CPT	6 months	Yes
5. Lee	19A	1LT	12 months	Yes
6. Bertram	**11A**	**1LT**	**18 months**	**Yes**
7. Williams	12A	MAJ	12 months	Yes
8. Johnson	17A	CPT	18 months	No
9. Ramirez	**11A**	**1LT**	**12 months**	**No**
10. Patel	25A	CPT	6 months	Yes
11. Jones	**11A**	**1LT**	**12 months**	**Yes**
12. Becker	12A	MAJ	6 months	No

2. Interruption task

From: MAJ Campbell
Subject: Deployable Soldiers
Do you have any 12A officers in Roster 1 who are not deployable? V/R, Maj Campbell

Please select your response: YES/NO

Appendix D

Problem Solver

Measurement Details	
Constructs: Problem Solver	**Subdimensions:** Representations, Approaches

Source (ATAF) Definition:

Problem Solver: A capacity to choose between best practices and unorthodox approaches to reach a solution; accomplishes the task.

Operational Definition*:*

Problem Solver: Understands multiple approaches for representing and solving problems.

Definition Decomposition:

- "Understands" – the Problem Solver trait is composed of a complex of knowledge and skills
- "Multiple approaches" – the Problem Solver trait is about breadth. An expert in one domain may proficiently solve problems for that domain, but not score highly on the Problem Solver trait because they lack breadth in problem solving knowledge and skills.
- "Representing and solving problems" – the Problem Solver trait is grounded in a basic information processing view of problem solving as a multi-staged process
    - A concise model of problem solving is summarized in four stages: a) Form initial representation, b) use methods to for a plan, c) reformulate the plan as needed, d) execute the plan and test the outcome (Glass & Holyoak, 1986)
    - The two initial stages are measured as two dimensions of the Problem Solver trait; these dimensions are Representations and Approaches.
    - Definitions and example items are below
    - Both dimensions are given equal weight in the Problem Solver score

Representations: Knowledge of and skill in forming external representations of an ill-defined problem

Approaches: Knowledge of and skill in identifying methods and plans for solving problems

Decisions:

- Focus on the "core" of problem solving
- Avoid creative problem solving – it is being measured elsewhere

Measurement Overview: A multiple-choice test of general problem-solving knowledge and skills, consisting of items measuring both dimensions. Scenario-based stems will be used to frame the questions

Measure Details:

Description:

- Representation Items: Test takers are asked to identify the best external representation of a problem given a goal (e.g., reframing, communicating, understanding the problem). Correct answers are verifiably correct based on prescriptions in the literature for creating external representations and visualizations.
- Approaches Items: Test takers are asked to identify the best approach for either solving or making progress toward solving a problem. Approaches items may also deal with formulating a plan based on chosen methods for solving a problem. Correct answers are determined based on both prescriptions in the literature and SME consensus.

Scoring:

Each item is scored as correct or not correct and total scores are based on the average across all items.

This assessment could be administered in CAT framework to reduce item exposure and testing time, while increasing the precision of the ability estimate.

Example Items:

Representations:

Each week, a person of interest goes out to a rendezvous point, meets with a contact, then returns home the next day. The person of interest always follows the same path to and from the rendezvous point and the trips always take the same amount of time. , You want to determine if there is a place along the path that the person reaches at the same time on both the trip out and the trip back, so you can form a plan for intercepting information from the person's mobile phone.

1. What is the best way to show this problem to come to a quick solution?
   A. With a short, descriptive paragraph laying out all the considerations in the scenario
   B. With a pair of overlapping circles representing the trip out and the trip back
   C. With a table of locations along the path and the times those locations are typically reached
   D. With a pair of overlapping line graphs showing location on path by time of day

Approaches:

2. A new system is malfunctioning without an apparent cause. All the following approaches are valid ways to solve a problem, but which one would work best in the above scenario?
   A. Work backward from the goal to determine the issue
   B. Think of an analogy or metaphor for the problem
   C. Leave the problem alone for a while, then come back to it later
   D. Guess at what the problem is and run a test to confirm

Appendix E

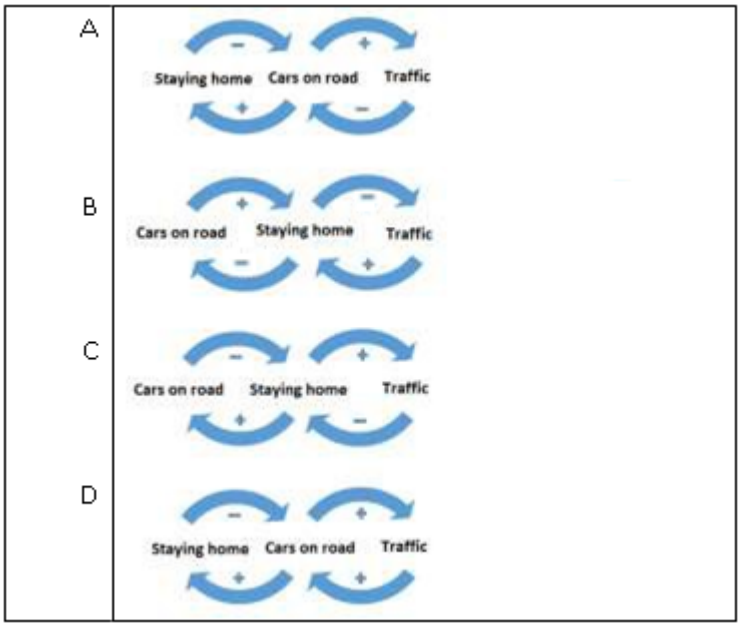
Systems Thinking

Measurement Details	
Construct: Systems Thinking	**Subdimensions:** None
Source (ATAF) Definition: Conceptualizes and understands relationships and arrangements within and between relevant components and structures	
Operational Definition*:* Conceptualizes and understands relationships and arrangements within and between relevant components and structures	
Decisions: • Will not use Systems Thinking Ability (STA) video game assessment as stimuli for the measure • Prefer to use a measurement strategy that minimizes technology requirements	
Measurement Overview: Test taker will respond to multiple choice items that describe the relationships between/among elements of a system and demonstrate their understanding of the relationships and dynamics in the system. Test taker will view a static image/graphic or read a text description that describes a fictitious system, then select the image/graphic or text description that corresponds to the system. System characteristics will be modeled after a variety of system types (i.e., biological, mechanical, social, etc.), but the assessment will be designed such that existing knowledge of specific systems is not required.	
Measure Details: *Description:* • Item Type 1: Test taker will read a text description of a system and select the best graphical representation of the relationships in that system • Item Type 2: Test taker will be presented with a graphic model or image of a complex system and be asked to select the multiple-choice response option that best describes the graphical system *Scoring:* Scores are based on total correct responses. This scoring strategy could be administered in CAT framework to reduce item exposure and testing time while increasing the precision of the ability estimate.	

Example Items:

1. Traffic on the highways around national holidays can be quite an obstacle. Over the decades, the traffic has become heavier with an increased number of cars on the roads. It is now to the point where many people do not bother to travel on holidays anymore and just stay home. Which of the following diagrams best represents this situation?

Response Options:



Correct Response: B

2. The model below shows relationships among three components. Which of the descriptions listed below best represents the model?

Response Options:

a. The more LEN there is, the more ACE there is likely to be. DRG decreases the release of ACE and LEN, so when there is more DRG, there is likely to be less ACE and LEN.

b. LEN causes the release of ACE. ACE causes the release of DRG. DRG causes the release of LEN.
c. The less LEN there is, the more ACE there is likely to be. DRG increases the release of ACE and LEN, so when there is more DRG, there is likely to be more ACE and LEN.
d. LEN decreases the release of ACE. ACE increases the release of DRG. DRG increases the release of LEN.

Correct Response: A

Appendix F

Perceptive

Measurement Details	
Construct: Perceptive	**Subdimensions:** None

Source (ATAF) Definition:

Effectively discerns the deeper meaning or significance of one's observations (e.g., events, people, and communication)

Operational Definition:

Effectively discerns the deeper meaning or significance of one's observations (e.g., events, people, and communication)

Decisions:

- Create a measure that is not focused specifically on social cognition
- Focus on the cognitive aspect of construct

Measurement Overview: Test takers are provided with written context and then shown a scene and asked to discern meaning from the scene. Test takers are asked to respond to two types of items: (1) items related to tasks identified in the scenario, and (2) items related to social aspects of the scenario. Test takers use a 5-point Likert scale to rate the likelihood of various statements or respond to a multiple-choice question.

Measure Details:

Description:

Item stimuli include a written item stem and static image of a related scene. After viewing the image for 10 seconds, test takers must respond to two types of questions:

- Item Type 1: Test takers must rate the likelihood of various statements related to the scenario and scene; Responses use a 5-point Likert scale
- Item Type 2: Multiple choice question for which test takers select the best response

Scoring:

Item developers will have the "best" answer in mind when writing the items. However, due to the potential lack of one, and only one, correct answer, an empirical scoring key and rational scoring will be considered.

Empirical Keying

Empirical would involve the following steps:

6. Split criterion-validation sample into a validation sample (80% of all respondents) and a holdout sample (20% of all respondents).

7. Compute average test taker scores on criterion of interest (e.g., Officer job/training performance).
8. Score response options such that response options associated with the highest criterion score (i.e., respondents who endorsed this response option had the highest criterion score when compared to those who endorsed other response options) are given a value of 1. The response option with the lowest criterion score would be given a value of 0, and any response(s) associated with a criterion score between the highest and lowest would be given a value of .67, .5 or .33 (depending on the number of responses that fell between the highest and lowest scores).
9. Compute biserial correlations between the rescored items and the criterion scores for both the validation and hold out samples.
10. Items that meet a certain threshold in both samples will be retained.

Rational Scoring

An alternative approach would be to use a rational based scoring key in which response options are assigned a value based on input by the test developers and SMEs. In this case, the response option determined to be the most plausible option would be given 1 point, the least plausible option which would receive 0 points and the other response options would be assigned partial credit according to how relatively plausible they are considered.

Example Items:

Note: Correct response options are bolded.

In this section, there will be a short scenario to read, followed by a photo image. After you review the photo, you will be asked a few questions about the situation.

On the drive to your job as the night manager at a local restaurant in Texas, you are thinking about what you need to do when you arrive. Here is the scene as you walk through the restaurant:

Source: Lang White, Tyle Morning Telegraph

1. Friday is usually busy during the day so you will need to take inventory and decide whether you should reorder anything for the next day's delivery. Common things you need to reorder are beer, pizza, french fries, and toilet paper. Rate the likelihood of each of the following statements (1=not at all likely to 5=highly likely):

Based on what you saw in the photo, how likely is it that….

1a. You will need to reorder beer?
1b. You will need to reorder pizza?
1c. You will need to reorder toilet paper?

2. As you walked through the restaurant, you noticed that everyone seemed to be watching something. Rate the likelihood of each of the following statements (1=not at all likely to 5=highly likely):

Based on what you saw, how likely is it that….
2a. They are watching the opening ceremony of the Olympics.
2b. They are watching a college soccer game.
2c. They are watching an international soccer game.
2d. They are part of a team or social club.

3. The daytime manager stops by on her way out to coordinate with you, and she says the current customers have been there for a few hours watching an Olympic soccer game. She asks you for an update on the game.

Based on what you saw, select the best response:

- **The US team scored a goal**
- The game just started
- The game is at halftime
- The opposing team scored a goal

4. How would you describe the mood in the room as you walked through?

Based on what you saw, select the best response:

- Very excited
- **Somewhat excited**
- Very mixed
- Anxious
- Very anxious

Response Rationale:
Question 1.
1a=1 or 2*;* Most of the people in the photo appear under 21 and no one is noticeably drinking beer.
1b=3 or 4; People in the photo are eating and it looks like they are eating pizza.
1c=5; There are a lot of people at the restaurant, so it is likely they are using the bathroom/toilet paper.

Question 2.
2a. 1 or 2 – people are cheering/on edge as if watching a competition
2b. 3? – people seem college-aged, so it is possible, but they are not wearing college team fan attire
2c. 4 or 5 – people seem to be dressed in red, white, and blue suggesting they are rooting for a US team
2d. 3 or 4? – it does seem like the people know each other but it is not necessarily clear that they are in a team or club