Technical Report 1361

Development and Preliminary Validation of the Strategic Thinking Mindset Test (STMT)

William S. Weyhrauch
Consortium Research Fellows Program,
Consortium of Universities of the Washington Metropolitan Area

June 2017

United States Army Research Institute for the Behavioral and Social Sciences

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

Department of the Army
Deputy Chief of Staff, G1

Authorized and approved:

MICHELLE SAMS, Ph.D.
Director

Research accomplished in-house for the
Department of the Army

Technical Review by

Kimberly A. Metcalf, U.S. Army Research Institute
Anna Sackett, U.S. Army Research Institute

REPORT DOCUMENTATION PAGE

1. REPORT DATE (dd-mm-yy)	2. REPORT TYPE	3. DATES COVERED (from. . . to)
June 2017	Final	Sept 2012 – May 2016

4. TITLE AND SUBTITLE	5a. CONTRACT OR GRANT NUMBER
Development and Preliminary Validation of the Strategic Thinking Mindset Test (STMT)	5b. PROGRAM ELEMENT NUMBER 622785
6. AUTHOR(S) William S. Weyhrauch (Consortium Post-Doctoral Fellow)	5c. PROJECT NUMBER A790
	5d. TASK NUMBER
	5e. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Consortium Research Fellows Program Consortium of Universities of the Washington Metropolitan Area 4214 King Street, First Floor Alexandria, VA 22302-1555	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) U. S. Army Research Institute for the Behavioral and Social Sciences ATTN: DAPE-ARI-RK 6000 6th Street Fort Belvoir, VA 22060-5610	10. MONITOR ACRONYM ARI
	11. MONITOR REPORT NUMBER Technical Report 1361

12. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited.

13. SUPPLEMENTARY NOTES

ARI Research POC: Angela Karrasch, FLRU (angela.i.karrasch.civ@mail.mil); Subject Matter POC: William S. Weyhrauch

14. ABSTRACT *(Maximum 200 words)*:

Senior Army leaders face complex problems. These problems lack a right answer, they feature interdependent variables, stakeholders with conflicting goals, and actors that adapt to any action taken. The Army needs to do everything possible to prepare Army leaders to meet these challenges of complexity. The Army must manage talent and develop strategic thinkers early and efficiently by providing the experiences, education, and opportunities for feedback that are necessary for skill development. This report describes the initial research on the strategic thinking mindset, a concept that can help the Army efficiently manage talent among its tactical leaders. The strategic thinking mindset is made up of three intellectual characteristics foundational to strategic thinking: flexibility, humility, and inclusiveness. A scenario-based assessment tool was developed to measure each of these characteristics in Company-grade officers. The Strategic Thinking Mindset Test (STMT) consists of 12 scoreable items useful for assessment, self-reflection, and group discussion. A pilot test of the STMT indicated evidence of construct validity and discriminant validity. Further research is needed to establish test-retest reliability and criterion-related validity, as well as study other methods for assessment and developing strategic thinking early in the leader development path. See also ARI Research Product 2018-02.

15. SUBJECT TERMS

strategic thinking, situational judgment test, intellectual flexibility, intellectual humility, intellectual inclusiveness, assessment, talent management

SECURITY CLASSIFICATION OF			19. LIMITATION OF ABSTRACT	20. NUMBER OF PAGES	21. RESPONSIBLE PERSON
16. REPORT Unclassified	17. ABSTRACT Unclassified	18. THIS PAGE Unclassified	Unlimited		Dr. Angela Karrasch Chief Fort Leavenworth Research Unit 913-684-9758

Standard Form 298

Technical Report 1361

Development and Preliminary Validation of the Strategic Thinking Mindset Test (STMT)

William S. Weyhrauch
Consortium Research Fellows Program,
Consortium of Universities of the Washington Metropolitan Area

Fort Leavenworth Research United States Army
Angela Karrasch, Chief

June 2017

DEVELOPMENT AND PRELIMINARY VALIDATION OF THE
STRATEGIC THINKING MINDSET TEST (STMT)

EXECUTIVE SUMMARY

Research Requirement:

Senior leaders in the Army manage extensive resources and make decisions that impact organizational strategic success in many ways. Strategic thinking is a critical requirement of these leaders. Talent management through selection and development relies on assessment, early and over time. Assessment can inform selection for positions and opportunities, but also coaching/mentoring, self-awareness, and self-development. Without early awareness and intentional development of a strategic thinking mindset, Army leaders face a difficult challenge in adopting this mindset later in their careers, after having succeeded at tactical and operational levels that prioritize a different mindset. Assessment tools are needed to identify Army leaders with a natural mindset for strategic thinking and to help develop the mindset in those who may lack the natural mindset, but have other strengths.

Procedure:

Using a situational judgment testing (SJT) methodology, the U.S. Army Research Institute for the Behavioral and Social Sciences (ARI) developed the Strategic Thinking Mindset Test (STMT). The STMT was designed to measure the extent to which a Company-grade Army officer approaches tactical-level problem scenarios with a mindset of intellectual flexibility, intellectual humility, and intellectual inclusiveness, characteristics that are foundational to strategic thinking.

Findings:

A 12-item SJT was designed over four stages of development. In the first stage, Soldiers provided key incidents from their experience of complex and ambiguous problems. In Stage 2, scenarios based on these incidents were presented to Soldiers for generation of response options based on high and low levels of flexibility, humility, and inclusiveness. In Stage 3, response options were evaluated and scored by Soldiers for effectiveness and representativeness of the construct. Finally, in Stage 4, the test was piloted and construct validity evidence was gathered.

The pilot test of the STMT revealed evidence of construct validity, but low internal consistency reliability. The test's three subscales (intellectual flexibility, inclusiveness, and humility) each correlated significantly with alternative measures of these characteristics, while scores on the test were distinct from general cognitive ability. The results suggested that a respondent's choice of what *not* to do in response to a scenario had greater construct validity than their choice of what they *would* most likely do.

Further research is needed to fully support the reliability and validity of the test. Specifically, an examination of test-retest reliability will better indicate reliability than internal

consistency, due to the complex nature of SJT items. Efforts to gather criterion-related validity evidence would also improve the use of the test.

Utilization and Dissemination of Findings:

The research findings can benefit several stakeholders in the Army. First, commanders can use the test in leader professional development settings to raise self-awareness and initiate discourse about how intellectual flexibility, humility, and inclusiveness manifest in tactical- and operational-level leadership and decision-making. Secondly, Army course instructors and curriculum designers can use the test to measure their students' mindset at the beginning and/or end of courses related to leadership and decision-making. Finally, Army researchers can use the test in further research studies on the career-long development of strategic thinking competencies to foster improved understanding of the benefits of various interventions and development opportunities as part of a talent management effort.

DEVELOPMENT AND PRELIMINARY VALIDATION OF THE STRATEGIC THINKING MINDSET TEST (STMT)

CONTENTS

Page

RESEARCH NEED AND CONCEPT DEVELOPMENT 1
Organization of the Report 2
Strategy 3
Strategic Thinking 4
Schools of Thought 4
Models of Strategic Thinking 5
Common Themes Among Strategic Thinking Models 7
Systems and Synthesis 9
Creativity 9
Directionality 9
Criticality 9
Awareness of Time 10
Adaptability and Opportunism 10
Breadth and Inclusion 10
Self-Awareness and Self-Control 10
Action Learning 10
Relation of Themes to Army Strategic Thinking Competencies 11
The Strategic Thinking Mindset 13
Characteristics of the Strategic Thinking Mindset 14
Flexibility 16
Inclusiveness 17
Humility 17
Other Examples of Mindsets 18
Strategic Thinking Mindset Change 19
The Strategic Thinking Mindset Test 20
Situational Judgment Test Format 20
Research Questions and Hypotheses 21

TEST DEVELOPMENT 22
Stage 1 Method 22
Sample 22
Materials and Procedure 24
Stage 1 Results 24
Stage 2 Method 25
Sample 25
Materials and Procedure 26
Stage 2 Results 27
Stage 3 Method 28
Sample 28
Materials and Procedure 29
Stage 3 Results 30

Stage 4 Method 35
Sample 35
Materials and Procedure 36
Stage 4 Results 37

DISCUSSION 46
Strengths of the Research 50
Limitations of the Research 51
Future Research 52
Implications 55

CONCLUSION 55

REFERENCES 57

APPENDIX A 66
APPENDIX B 70
APPENDIX C 78
APPENDIX D 83

LIST OF TABLES

TABLE 1. RELATION OF MODELS TO COMMON THEMES 8

TABLE 2. RELATION OF COMMON THEMES TO ARMY COMPETENCIES AND ENABLERS 12

TABLE 3. STAGE 1 SAMPLE DEMOGRAPHICS 23

TABLE 4. STAGE 2 SAMPLE DEMOGRAPHICS 25

TABLE 5. STAGE 3 SAMPLE DEMOGRAPHICS 29

TABLE 6. RESPONSE OPTION KEYING DATA—FLEXIBILITY ITEMS 31

TABLE 7. RESPONSE OPTION KEYING DATA—HUMILITY ITEMS 32

TABLE 8. RESPONSE OPTION KEYING DATA—INCLUSIVENESS ITEMS 34

TABLE 9. STAGE 4 SAMPLE DEMOGRAPHICS 35

TABLE 10. INTERITEM CORRELATION MATRIX (ALL ITEMS) 39

TABLE 11. ITEM-SCALE AND VALIDITY CORRELATIONS (ALL ITEMS) 40

TABLE 12. ITEM-SCALE AND VALIDITY CORRELATIONS (FINAL SCALE ITEMS)...41

TABLE 13. GROUP 1 SCALE VALIDITY CORRELATIONS (FINAL SCALE ITEMS) 42

TABLE 14. GROUP 2 SCALE VALIDITY CORRELATIONS (FINAL SCALE ITEMS) 43

TABLE 15. ITEM-SCALE REJECTION-ENDORSEMENT CORRELATIONS (FINAL SCALE ITEMS) 44

TABLE 16. SCALE LEVEL CORRELATIONS BETWEEN REJECTION AND ENDORSEMENT SCORES (FINAL SCALE ITEMS).....45

LIST OF FIGURES

FIGURE 1. DIRECT COMPARISON OF LIEDTKA AND YARGER MODELS 5

FIGURE 2. RELATION OF MINDSET CHARACTERISTICS TO COMMON THEMES 15

DEVELOPMENT AND PRELIMINARY VALIDATION OF THE STRATEGIC THINKING MINDSET TEST (STMT)

Research Need and Concept Development

Over the past several decades, organizations of all types, including the U.S. Army, have evolved their structures and processes in response to increased complexity and environmental unpredictability (Alberts & Hayes, 2003; Huber, 2011). Common organizational structures in the past were highly bureaucratic, characterized by clear hierarchical levels and functional divisions. In contrast, many organizations today are under pressure to become increasingly agile and responsive to the environment, with many developing structures and processes that are more decentralized and flexible (Huber, 2004). The Army doctrine of mission command, as described in a white paper by former Chairman of the Joint Chiefs of Staff General Martin Dempsey (Dempsey, 2012), promotes the philosophy of countering environmental uncertainty with centralized planning and decentralized execution, with commanders building trust and shared vision in their unit and allowing subordinates to execute within the commander's intent.

In both the private and public sectors, organizations must operate within complex adaptive systems. These systems feature a large number of individual agents continuously interacting and adapting to each other uniquely, producing system-wide effects that are largely unpredictable and constantly evolving (Anderson, 1999). The environmental context of an organization is critical to determining how best to go about conducting strategic planning. Stieglitz, Knudsen, and Becker (2015) describe how the importance of strategic exploration and flexibility vary depending on the environment's dynamism. Environments characterized by persistent trends and few structural shifts require different strategic choices than those characterized by volatility, uncertainty, complexity, and ambiguity (VUCA).

When an organization exists within a VUCA environment, as do the Army and U.S. government (Gerras, 2010; Jacobs, 2002), executing strategy to ensure success becomes increasingly difficult and requires a greater investment in developing competence in strategic thinking throughout the organization, particularly among its leaders. As such, strategic thinking has become an imperative for modern businesses (Duhaime, Stimpert, & Chesley, 2012) and governments (Yarger, 2008). Sustained research on strategic thinking and its development is needed to ensure long-term organizational success and to overcome what may be a weakness in many organizations (Bethel, Prupas, Ruby, & Smith, 2010; Krepinevich & Watts, 2009; Sackett, Karrasch, Weyhrauch, & Goldman, 2016).

Research on strategic thinking has focused not only on the conceptualization of what strategic thinking entails (Graetz, 2002; Liedtka, 1998a), but also how it is developed (Bonn, 2001, 2005; Goldman & Casey 2010; Eifler, 2012). There is a consensus that time and experience are needed to develop strategic thinking (Dragoni, Oh, Vankatwyk, & Tesluk, 2011; Goldman, 2008). Goldman (2008) described the importance of spearheading a growth initiative or doing strategic planning activities in order to develop strategic thinking. The Army Research Institute for the Behavioral and Social Sciences (ARI) has recently conducted a line of research exploring the needs for better strategic thinking development in the U.S. Army (e.g., Wolters, Grome, & Hinds, 2013; Sackett et al., 2016). The Army Leader Development Strategy (U.S.

Army Training and Doctrine Command, 2009) emphasizes talent management and the need for broadening assignments that give officers the perspective and knowledge necessary to be effective at strategic levels of command. An ongoing Army effort, partially informed by this and other ARI research, is being led by the Strategic Education Sub-Committee of the Army Learning Coordination Council to improve the education of strategic thinkers across Professional Military Education (PME) and the force. Ongoing research is needed to determine how best to guide this process to ensure that these efforts are effective.

If certain experiences are necessary to develop strategic thinking, a practical problem quickly becomes apparent: organizational decision makers must decide how to assign opportunities for such experiences. That is, not everyone can be given the developmental opportunities in the organization. In order to optimize the development of strategic thinkers for an organization, leaders must make informed decisions about who has *potential* for success at a higher level so that they can be properly mentored and placed in positions that will develop and hone their strategic thinking ability through exposure and practice. There are a variety of constructs that could be used to predict one's potential for strategic thinking. For example, general and technical knowledge, intelligence, creativity, adaptability, and personality have been proposed as predictors of strategic cognitive readiness in the military (Grier, 2012). A similar list from the management literature includes cognitive ability, personality, and work experience (Dragoni et al., 2011).

The purpose of the current research is to contribute to this literature on the understanding and development of strategic thinking by conceptualizing a new construct with the potential to predict and forecast strategic thinking ability: the strategic thinking mindset (STM). Furthermore, this research will aim to develop a measure of the strategic thinking mindset using real-world ambiguous problem scenarios. The strategic thinking mindset refers to *a tendency to approach problems in a manner that is consistent with the cognitive elements of strategic thinking, specifically intellectual flexibility, intellectual humility, and intellectual inclusiveness.*

Organization of the Report

This report is structured to provide a concise, but thorough, summary of the development of the strategic thinking mindset concept and test. The anticipated audiences are military leader development professionals and measurement researchers. For the sake of the latter audience, some extra content has been included in Appendix A. The major sections of this report are:

- **Strategy**: what strategy means, how it varies across disciplines, and how it is conceptualized in this research.
- **Strategic thinking**: what strategic thinking means, how it is characterized by different schools of thought and individual scholars, and common themes among them
- **Strategic thinking mindset**: linking the common themes of strategic thinking to three key fundamental characteristics of an individual's mindset
- **Test development Stage 1**: gathering of key incidents for test scenarios
- **Test development Stage 2**: refinement of scenarios and generation of response items
- **Test development Stage 3**: scoring and selection of response items

- **Test development Stage 4**: pilot testing and construct validation of test items
- **Discussion**: summary of the findings, strengths and weaknesses of the research, and implications for future research and practice

Strategy

The English word strategy is rooted in the Greek term *strategos* meaning *military leader*. The U.S. military defines strategy as "a prudent idea or set of ideas for employing the instruments of national power in a synchronized and integrated fashion to achieve theater, national, and/or multinational objectives" (U.S. Department of Defense, 2017, p. 224). Despite its origins in war, however, the term is now frequently employed in other contexts. For example, managerial scholars have defined strategy in terms of outmaneuvering competitors by "finding alternative ways of competing and providing customer value" (Abraham, 2005, p. 5). In game theory, strategy is a mathematical concept, specified by a functional equation which determines an action given the sequence of previous actions (Pelc & Pelc, 2009). In human cognitive development, strategies are "non-obligatory, goal-directed activities designed to enhance task performance that are potentially available to conscious awareness" (Schwenck, Bjorklund, & Schneider, 2009, p. 1034). Regardless of the specific discipline defining the concept, it is clear that a constant theme underlying strategy concepts is the notion of a carefully developed plan meant to guide the actions of a party in seeking an objective.

Three prominent scholars of strategy, Lawrence Freedman, Colin Gray, and Michael Porter, have differed somewhat in their definitions of strategy. In Freedman's (2013) work on summarizing the history of the use of strategy, he provides a broad definition of strategy as "the art of creating power" (2013, p. xii). Freedman structures his history according to three domains in which strategy has evolved: military strategies of force, social-political revolutions in which strategy occurs from the bottom up, and business management strategies in which strategy is determined from the top down. For Freedman, the master strategist, one who is able to foresee the future and execute a planned strategic victory, is a myth. The best a strategist can do is identify ways to improve the group's position in relation to strategic objectives, step by step, rather than through a pre-determined sequence of moves. Gray's discussion of strategy (see Gray, 2011; 2015) focuses more narrowly on the use of strategy by governments and their militaries. Gray defines military strategy as "the direction and use made of force and the threat of force for the purposes of policy as decided by politics" (Gray, 2015; p. 21) emphasizing the primacy of politics, which determine policies, and in turn establish strategic objectives. The idea of positioning is also central to strategy as conceptualized by Porter (1996). Strategic success in this context is determined by position relative to peer competitors. Porter emphasized that operational effectiveness is necessary, but not sufficient, for strategic success.

For the current research, strategy (and thus strategic thinking) should be understood in the same context as Freedman (2013), with a broader scope than the private sector, competitive consumer market context of Porter (1996), or the national security context of Gray (2015) and Yarger (2008). The purpose of strategy is the positioning of an organization effectively in a highly complex, dynamically adaptive environment. This context-generic perspective is perhaps best described as "future-oriented behavior concerned with [the] dynamic and complex relationship of the organization with its environment. It is a continuous process…which should

produce an improved future state for the organization" (Wheatley, Anthony, & Maddox, 1991; p. 52).

Strategic Thinking

Schools of thought. Extensive scholarship within the military, private sector, and other academic disciplines has been conducted on strategy development, strategic planning, and a wide assortment of related concepts. More recently, researchers have focused on strategic *thinking*, primarily at the individual level (e.g., Liedtka, 1998a), but also with a focus on how organizational culture inhibits or supports strategic thinking (e.g., Goldman & Casey, 2010; Bonn, 2001; 2005).

Disciplines vary in their definition of strategy. Likewise, definitions of strategic thinking vary. Yet, the differences are largely outweighed by commonality. Two broad schools of thought exist on how strategy is developed and what sort of thinking is required. These schools mirror the general question of whether strategy is deliberate or emergent (i.e., planned carefully in advance by specific people or something that happens over time through the actions and decisions of many).

The first school of thought suggests that strategy development and strategic thinking are primarily analytical activities that rely on deliberate processes to frame the environment and determine the most effective strategy. This can be described as the strategic planning school. An example of strategic thinking viewed from this perspective is that of Porter (1987), who argues for the importance of formal long-range planning systems to occasionally force managers to think beyond day-to-day concerns. For example, he proposed a five forces framework for analyzing a system (e.g., an industry) and developing strategies during long-range planning (Porter, 2008). Others in this school often advocate their own unique process approach to strategic thinking, such as the highly systematic six-step process of Zabriskie and Huellmantel (1991) or the similarly deliberate cognitive mapping approach of Eden (1990). SWOT analysis, whereby an organization's strengths, weaknesses, opportunities, and threats are studied, is another common example.

The strategic planning approach was updated by Mintzberg (1994a, 1994b), who described a new approach that challenged the focus on analysis by emphasizing synthesis and the need for strategy to emerge over time through learning. This second school of thought, the learning school approach, emphasized the difference between strategic *planning* and strategic *thinking*. According to Mintzberg, strategic planning tools are valuable in informing and implementing strategy, but can hinder strategic thinking if inflexible planning processes dominate strategy creation and limit creativity. Thus, these two primary schools of thought differ on the point of strategy being deliberate or emergent.

Mintzberg and contemporary scholars now tend to advocate positions that combine both schools, wherein strategic thinking consists primarily of continuous learning and adaptation, but is appropriately informed by deliberate analytical processes of strategic planning. This balanced view emphasizes strategic planning and strategic thinking as "distinct, but interrelated and complementary processes" (Heracleous, 1998, p. 482).

Models of strategic thinking. Currently, models of strategic thinking tend to align with the learning school with an emphasis on learning and emergent strategy. However, many models have expanded strategic thinking to include both the synthetic elements from the learning school and the deliberate, intentional elements of the planning school. Several of the most notable conceptual models of strategic thinking, as identified from military and management scholars, are reviewed below. From these, the common themes underlying strategic thinking became clear. This multidisciplinary understanding of strategic thinking provided the basis for identifying the characteristics of a strategic thinking mindset. In regard to the exact nature of each author's concept, some authors cited here may prefer a particular descriptor (e.g., framework, competencies, factors, etc.). The focus of this section is on the conceptual constructs referenced by each author. For the sake of clarity and consistency, the words *model* and *element* are used throughout. To make the process of model comparison and the synthesis of common themes clearer, two prominent models are introduced and compared directly. Beyond these two, the models will be presented individually. However, the same process of comparison and synthesis was conducted with all models in mind to identify the common themes that follow.

A well-known management model described by Liedtka (1998a) and a prominent military-focused model from Yarger (2008) are directly compared below. As shown in Figure 1, these two models clearly align in certain ways, e.g., the focus on systems and thinking in time. Yarger's (2008) inclusion of ethical thinking and Liedtka's (1998a) inclusion of intent-focus, however, are unique aspects. Yarger's critical and creative thinking are both related to Liedtka's intelligent opportunism and hypothesis-driven elements (e.g., to generate a hypothesis is a creative act; critical thinking is required when evaluating the hypothesis in relation to available data; opportunism implies recognition of an opportunity requiring critical thinking and creative thinking in developing a new approach). Mapping each model of strategic thinking onto all others in this way is unnecessary. However, this is the process by which the common themes across disparate models were identified. It is clear from this example that the different models need not be seen as mutually exclusive, but complementary and supplementary to each other.

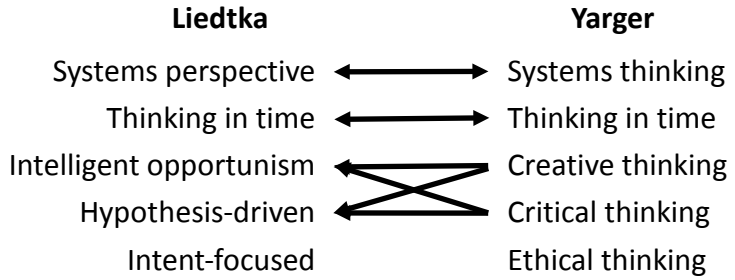

Figure 1. Direct comparison of Liedtka and Yarger models.

A model described by Bonn (2001) includes elements of (a) holistic understanding, (b) creativity, and (c) vision for the future. Bonn's work (2001; 2005) has been widely cited, along with Liedtka (1998a; 1998b). Heracleous (1998) is also commonly cited in distinguishing strategic thinking and strategic planning, comparing strategic thinking to double-loop learning (Argyris, 1991). Double-loop learning involves the questioning of rules and assumptions governing decision making, as opposed to single-loop learning wherein critical thinking can take place, but does not question the broader framework by which options for a decision are

determined. The Heracleous model also includes synthetic, divergent, and creative thought processes as elements underlying strategic thinking.

More recently, Duhaime, Stimpert, and Chesley (2012) devoted an entire textbook to understanding the importance of strategic thinking in modern business. Their model, framed as the characteristics of effective strategic thinkers, includes the following elements:

- "experts … [who] link disparate strands of information and … consider a broad array of scenarios and outcomes" (p. 69)
- "able to think dynamically … and be able to anticipate the future" (p. 70)
- "managers (and their firms and businesses) must be good learners" (p. 70)

The models depicted thus far focus on characteristics of thinking that epitomize strategic thinking. A model from Casey and Goldman (2010) identified four continuous and iterative activities of strategic thinking, based on a model of strategic thinking with four elements:

Activities

- scanning
- questioning
- conceptualizing
- testing

Elements

- conceptual
- systems-oriented
- directional
- opportunistic

Graetz (2002) also attempted to distinguish strategic thinking from strategic planning, proposing five attributes for strategic thinking and five for strategic planning. Graetz emphasized that both deliberate (planning) and emergent (thinking) approaches to strategy making must coexist, despite the differing attributes.

Strategic Thinking

- synthetic
- divergent
- creative
- intuitive
- innovative

Strategic Planning

- logical
- systematic
- conventional
- prescriptive
- convergent

A model from the U.S. Army's War College, as described by Waters (2011), advocates a balanced approach in which strategy making is an art and science. The Waters model includes a diverse array of elements and activities, including:

- critical thinking
- thinking in time
- synthesis
- systems thinking
- creative thinking
- futuring
- divergent and convergent thinking
- environmental scanning
- judgment of risk and reward
- strategic thinking foundation
  - self-awareness of one's biases and assumptions
  - openness to discourse and reflection

Another model from a military scholar includes four elements: (a) systems thinking, (b) visioning, (c) scanning the environment, and (d) scenario planning (McCauley, 2012). This model combines characteristics and types of thinking, as with Liedtka (1998a) and Yarger (2008), respectively, but also thinking activities, similar to Casey and Goldman (2010). While not explicitly discussing strategic thinking, Salmoni, Hart, McPherson, and Winn (2010) summarize the cognitive characteristics needed of military strategic leaders in the future, including (a) emphasis on how to think, rather than what to think (i.e., metacognition), (b) flexibility and openness to many disciplines, and (c) tolerance of iterative problem solving and lacking perfect solutions the first time. Finally, Yorks and Nicolaides (2012) described a model for how to develop mindsets for strategic insight. They define strategic insight as resulting from (a) engaging with diverse perspectives, (b) assessing trends in divergent domains, (c) making assumptions explicit, and (d) challenging those assumptions.

Some models of strategic thinking are embedded in measurement approaches. The Leader's Strategic Mindset assessment (Pisapia et al., 2005) was developed around three elements: (a) systems thinking, (b) reframing, and (c) reflecting. Dragoni and colleagues (2011) conducted an empirical study of the antecedents of strategic thinking competency. The assessment center criteria used as outcome variables were operationalized as (a) articulating a vision, (b) demonstrating sound business judgment, and (c) attending to global business issues. Finally, a report by Grier (2012) discussed what elements should be considered in assessing military cognitive readiness at operational and strategic levels of command:

- general knowledge and abilities
- cognitive capabilities
- creativity
- adaptability
- certain personality traits
  - hardiness
  - self-control
  - need for certainty

Common themes among strategic thinking models. The purpose of reviewing and comparing the many disparate competency models was to identify the shared themes underlying the concept. In integrating and synthesizing the models into a broader understanding of strategic thinking, nine themes clearly emerged:

- Systems and synthesis
- Creativity
- Directionality
- Criticality
- Awareness of time
- Adaptability and opportunism
- Breadth and inclusion
- Self-awareness and self-control
- Action learning

The themes are presented below, in descending order of their prevalence across the models. These themes provide the basis for the current concept of strategic thinking mindset. A summary of the relation of each model to the common themes is provided in Table 1.

Table 1

Relation of Models to Common Themes

	Systems & synthesis	Creativity	Direction-ality	Criticality	Aware-ness of time	Adaptability & opportunism	Breadth & inclusion	Self-awareness & self-control	Action learning
Heracleous (1998)	X	X		X					
Liedtka (1998b)	X		X		X	X			X
Bonn (2001)	X	X	X						
Graetz (2002)	X	X							
Pisapia et al. (2005)	X	X		X		X			
Yarger (2008)	X	X		X	X				
Casey & Goldman (2010)	X	X	X	X		X			X
Salmoni et al. (2010)						X	X	X	
Dragoni et al. (2011)	X		X				X		
Waters (2011)	X	X	X	X	X			X	
Duhaime et al. (2012)					X	X	X		X
Grier (2012)		X				X		X	
McCauley (2012)	X		X		X		X		
Yorks & Nicolaides (2012)				X	X		X	X	

Systems and synthesis. The most pervasive theme present in the theoretical models reviewed was that of systems and synthesis. This theme refers to the process of gaining new awareness by combining parts to perceive a whole with unique properties emerging from the combination of parts. This theme was present in 10 of the 14 models summarized and includes elements such as systems orientation (Casey & Goldman, 2010), synthetic thinking (Heracleous, 1998), holistic understanding (Bonn, 2001), and attending to global issues (Dragoni et al., 2011).

Systems thinking and the act of synthesis emphasize a higher-level collection of parts interacting and combining with each other to produce a collective effect or pattern. For most organizations, the primary system of interest would be the organization itself. However, systems thinking also requires the consideration of other systems in the environment, particularly those that subsume the organization and exist within it. Furthermore, the characteristics of an environment play a role in systems thinking. Synthesis tends to focus more on the combination of processes and concepts when contrasted with systems thinking, which tends to emphasize the interaction of tangible entities.

Creativity. The second integrating theme, creativity, was readily apparent in eight of the models. The creativity theme subsumes concepts such as innovative thinking (Graetz, 2002), reframing (Pisapia et al., 2005), and conceptual thinking (Casey & Goldman, 2010). For Casey and Goldman, conceptual thinking refers to theorizing new abstract ideas, which are then tested in the strategic environment. The references to creativity throughout the models were grounded in the creation of something new, whether it was a new process, technique, idea, or narrative of a problem. Creativity and synthesis are closely related concepts; therefore, these themes overlap to some degree. Synthesis was placed alongside systems as its own theme, rather than under this theme, due to its added emphasis on the joining together of existing system or organizational elements for added value. The creativity theme was meant to describe the generation of an idea from a less tangible source, such as subconscious processing or a connection made with something previously thought to be irrelevant to the organization.

Directionality. Directionality, the third theme, consists of a dedicated focus on seeking a desired future condition for the organization. Present in six of the models, this theme should be considered synonymous with concepts such as intent-driven thinking (Liedtka, 1998a) or vision for the future (Bonn, 2001). Directionality was an element clearly seen in almost every definition of strategy. Military organizations often refer to this as seeking an end-state; organizations encapsulate it in a vision statement. Whether the strategy is being developed at a geo-political level, to win a war, corner a market, succeed in a game, or merely find satisfaction in life, there is always a goal in mind, a point toward which thinking is oriented. An aspect of strategic thinking would include determining what the point of direction is or should be, but also how best to get there.

Criticality. Criticality was present in six of the models, including elements such as questioning (Casey & Goldman, 2010), challenging assumptions (Yorks & Nicolaides, 2012), and double-loop learning (Heracleous, 1998). This theme addressed the importance of challenging and questioning ideas or assumptions, as a means of affirming or disconfirming their validity. For an argument or assumption to be critically evaluated does not necessarily require it

to be fundamentally changed or abandoned. Rather, criticality is about explicitly acknowledging and evaluating ideas on their merits.

Awareness of time. The next integrating theme was awareness of time, a concept that may be less intuitive, but no less important, as an aspect of strategic thinking. Included in six models, this theme incorporates elements such as thinking in time (Yarger, 2008), anticipating the future (Duhaime et al., 2012), futuring (Waters, 2011), and assessing trends (Yorks & Nicolaides, 2012). Awareness of time refers to the consideration of how an issue is situated in past events, present contexts, and anticipated or potential futures. Having an awareness of time is highly important in dealing with complex adaptive systems, as the passing of time is inherently associated with changes in that system. With changes in any system, comes the need to check and adapt one's understanding and approach in dealing with problems in that system.

Adaptability and opportunism. The next theme, adaptability and opportunism, refers to changing one's approach or creating a new approach when key conditions in the environment change or are revealed to be different than was thought. This theme was present in six of the models summarized, based in elements such as adaptability (Grier, 2012), thinking dynamically (Duhaime et al., 2012), and intelligent opportunism (Liedtka, 1998a). The learning school of strategy, discussed at length by Mintzberg (1994a) and Senge (1990), is centered almost entirely on this point. Complex adaptive systems are so ambiguous and dynamic that it is essentially impossible to fully understand them and predict what will happen with certainty. Therefore, strategy-making must include room for adapting to unforeseen problems or taking advantage of unexpected opportunities.

Breadth and inclusion. The next theme, breadth and inclusion, refers to an openness to and equitable consideration of many diverse viewpoints. Present in five of the models, this theme was based on both the intentional inclusion of many viewpoints and disciplines (e.g., engaging with diverse perspectives, Yorks & Nicolaides, 2012), but also comprehensively searching the environment for any knowledge that may have some bearing on a problem (e.g., environmental scanning, McCauley, 2012).

Self-awareness and self-control. The theme of self-awareness and self-control was present in four of the models. This theme overlaps with that of criticality (particularly the aspect of self-criticism), but is more specifically about one's willingness and ability to maintain self-awareness, not only of weaknesses, but also of basic assumptions. Good strategic thinkers understand how their thinking is limited and intentionally counteract those limitations. This theme includes elements such as metacognition (Salmoni et al., 2010) and making assumptions explicit (Yorks & Nicolaides, 2012).

Action learning. The final theme, action learning, links closely to the adaptability and opportunism theme, as learning takes place as new information is incorporated, creating or changing one's knowledge of a strategic situation. However, the action learning theme covers the extension of this idea seen in three of the models: Liedtka's (1998a) concept of hypothesis-driven, Duhaime et al.'s (2012) organizational learning, and Casey and Goldman's (2010) concept of testing. These concepts emphasize the role of generating strategies by continuously developing concepts about the environment, implementing them, studying their impact on the environment, and using the results to learn and improve the strategy. In short, action learning is

about putting a plan or idea in action and evaluating the results in order to revise the plan or idea. This is the essence of the scientific method: that of developing a hypothesis, testing it, and interpreting the results as evidence confirming or disconfirming the hypothesis. In a sense, strategic thinkers must be applied scientists in their strategic environment.

Relation of themes to Army strategic thinking competencies. A recent research effort examining the U.S. Army's capability gaps regarding strategic thinking examined several relevant issues, such as formal development, assignments, and competencies (Sackett et al., 2016). A part of this effort included the creation of a competency model for strategic thinking, consisting of six core competencies and four enabling competencies. The core competencies were meant to describe the actual cognitive process of strategic thinking, contrasted by the enabling competencies, that describe skills needed to translate the individual strategic thinking process into an impact on strategic actions.

Core competencies

- comprehensive information gathering
- learning
- critical thinking
- innovative thinking
- thinking in time
- systems thinking

Enabling competencies

- knowledge
- collaboration
- communication
- emotional regulation

In Table 2, these core and enabling competencies are presented in relation to the nine common themes identified in this research. As with the models described above, there is a great deal of overlap among the constructs, despite some differences in terminology and organization of the constructs. Table 2 shows how the fundamental constructs line up with each other, determined through a close examination of each element's conceptualization. In brief, the Army competency model has a wider scope of focus. That model includes the notion of enablers that allow the individual strategic thinker to effectively influence the procedures of Army strategy. Several of the themes align directly with a core or enabling competency. For example, the theme of criticality and the competency of critical thinking. In some cases, multiple themes align with a single broader competency. For example, the themes of action learning and adaptability and opportunism both fall under the broader umbrella of the learning competency. Conversely, the breadth and inclusion theme primarily aligns with the core competency of comprehensive information gathering, but also relates closely to the enabling competency of knowledge, the latter referring to the raw material (in the form of explicit and tacit knowledge) that is processed during strategic thinking.

Table 2

Relation of Common Themes to Army Competencies and Enablers

	Core competencies							**Enablers**		
	Comprehensive information gathering	Systems thinking	Critical thinking	Learning	Innovative thinking	Thinking in time	Knowledge	Emotional regulation	Collab-oration	Commun-ication.
Breadth & inclusion	X						X			
Systems & synthesis		X								
Criticality			X							
Action learning				X						
Adaptability & opportunism				X						
Creativity					X					
Awareness of time						X				
Directionality						X				
Self-awareness & self-control								X		

The two enabling competencies that reference interactions with others (collaboration and communication) do not align with a theme, because most models of strategic thinking tend to focus on the individual's cognitive process. That said, there is value to the Army in understanding the enablers that allow strategic thinkers to execute actions based on their strategy.

The Strategic Thinking Mindset

Successful strategic thinking certainly requires a high level of cognitive ability. Fluid and crystallized intelligence, memory, pattern recognition, and rapid information processing are likely all important aspects of being able to understand a strategic environment well enough to develop and implement an effective strategy. However, it is also clear from the models and themes described above that cognitive ability alone is not sufficient for successful strategic thinking. Other intellectual characteristics are also required that may or may not be held by those with the greatest intelligence. These characteristics constitute the *strategic thinking mindset*.

The distinction between the *ability* related to a thinking competency (strategic thinking) and the *mindset* for the same competency is a subtle, but crucial one. As depicted by Yorks and Nicolaides (2012):

Generative Strategic Insight = *f*(Competency x Capacity).

In other words, the generation of strategic insights is a function of the combination of one's competency for strategic insights and capacity. In this framework, competency refers to one's ability (e.g., reasoning, knowledge, cognitive tools). Capacity refers to the fitness of one's mindset or viewpoint. To state this more simply, effective strategy development requires both the proper cognitive ability tools for strategic thinking (which most of the strategic thinking literature focuses on) and the proper mindset, the focus of the current research.

The strategic thinking mindset is conceptually distinct from cognitive ability. Having a strategic thinking mindset reflects a viewpoint and approach to understanding problems that is consistent with the requirements of strategic thinking. This viewpoint is distinct from the information processing involved in the execution of strategic thinking. However, the level of one's strategic thinking mindset would likely predict an individual's chances of being an effective strategic thinker in the future, alongside other cognitive capabilities and personality traits that are less malleable.

The distinction between an ability and the mindset for that ability is similar in some ways to emotional intelligence (EI; Van Rooy & Viswesvaran, 2004). Many have extolled the virtues of this alternative view of intelligence (Goleman, 1995; Mayer, Salovey, & Caruso, 2000), while others have criticized the lack of substantive scholarly work supporting the existence of the concept as defined and its proposed relationship to job performance (Zeidner, Matthews, & Roberts, 2004). The literature on EI coheres around two types of models: ability models, based on the work of Mayer et al. (2000), and mixed models, represented by the work of Goleman (1995). Ability models focus on the cognitive processing of emotional information, based on four competencies: identification, understanding, usage, and self-regulation (Salovey, Bedell, Detweiler, & Mayer, 2000). Conversely, mixed models are looser in structure and include a wider variety of competencies, such as motivation, temperament, and social skills. Cherniss and Goleman (2001) cite four core competencies in their mixed model of EI: self-awareness, self-regulation/management, social awareness, and relationship management/social skills.

The competencies associated with the mixed model of EI are more akin to the mindset concept, rather than the more strictly cognitive, information-processing perspective of the ability model. However, the components in these EI models are in some cases unrelated to strategic thinking. As noted by Zeidner et al. (2004), the question of "whether placing all such competencies under the EI banner confuses, rather than clarifies, the role of emotional competencies in the workplace would seem a contentious point" (pp. 378–379). Zeidner et al. further argue that "dealing with distinct but possibly interrelated competencies may be more tractable for research and practical purposes" (p. 379). Although understanding the emotional aspects of job performance is important, for the sake of clarity, the strategic thinking mindset is not conceptualized as a form of EI. Rather, it is a set of interrelated characteristics that form the foundation of an important work-related competency.

Characteristics of the strategic thinking mindset. The nine major themes common to strategic thinking models served as a basis upon which to conceptualize the characteristics of the strategic thinking mindset. The process of identifying these characteristics was one of synthesis among the themes, extracting their shared concepts, as well as examining the specific model elements that contribute to each theme and analyzing their conceptual definitions. As detailed below and shown in Figure 2, three intellectual characteristics underlie most of the common themes of strategic thinking: (a) intellectual flexibility, (b) intellectual inclusiveness, and (c) intellectual humility. These characteristics are malleable, in the sense that the characteristics can be developed over time with the proper focus, but not transient, meaning they are not so superficial as mood or states of emotion; rather, they are deeply ingrained in a person's cognitive behavioral habits.

The ideal strategic thinker would have a mindset consistent with these characteristics, along with substantial cognitive ability, and the knowledge and skills developed through experience, education, and training in the field. As discussed in the sections that follow, each of these characteristics could be taken beyond a certain threshold and begin to hinder performance, depending on the nature of the job (Le, Oh, Robbins, Ilies, & Holland, 2011). The characteristics are intellectual in nature, characteristics of a person's thought rather than merely characteristics of that person's behavior. In other words, the apparent expression of these traits in behavior, without truly thinking and feeling in these ways, would not constitute a strategic thinking mindset.

In the sections that follow, each of the three characteristics that comprise the strategic thinking mindset are described, with particular emphasis on how each characteristic contributes to strategic thinking capability. Note that the characteristics do not comprehensively underlie or influence all of the strategic thinking themes. For example, as indicated in Figure 2, there is no obvious link between the strategic thinking mindset characteristics and the strategic thinking themes of directionality (i.e., a focus on seeking a desired future condition for the organization) or awareness of time (i.e., the consideration of how an issue is situated in past, present, and future contexts). Finally, the strategic thinking mindset is surely only one way to determine potential for strategic thinking, there are likely other antecedents that can predict strategic thinking ability.

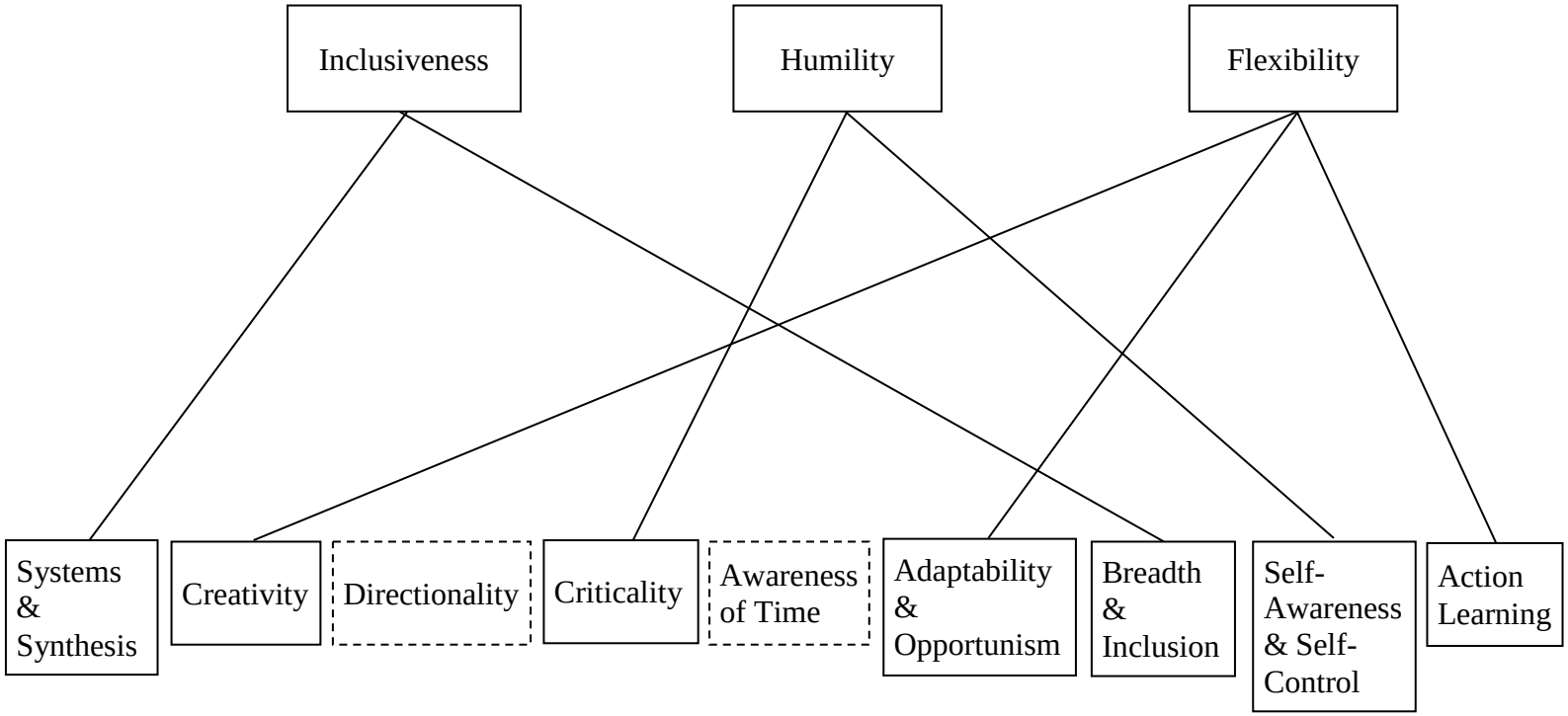

Figure 2. Relation of mindset characteristics to common themes.

Two of the identified themes are not linked to a mindset characteristic, these are shown with dashed lines.

Flexibility. The first mindset characteristic is flexibility. A mindset featuring intellectual flexibility is characterized by a willingness and proclivity to adjust one's understanding, opinions, or approach when conditions change or new information is presented. In essence, flexibility means not resisting necessary or optimal change. The psychological bias of escalating commitment (also referred to as sunk cost fallacy) refers to a tendency to continue with a course of action when that action is no longer rational (Sleesman, Conlon, McNamara, & Miles, 2012; Staw, 1976). Having a flexible mindset is crucial to the ability to avoid making errors of escalating commitment in the execution of strategy.

The importance of a mindset favoring flexibility was seen in three of the common strategic thinking themes: (a) adaptability and opportunism, (b) action learning, and (c) creativity. Being willing and comfortable to deal with change, major or minor in nature, is crucial to being adaptive and taking advantage of opportunities. Specific model elements of reframing (Pisapia et al., 2005), flexibility (Salmoni et al., 2010), and adaptability (Grier, 2012) clearly referenced the ability to change or adjust when necessary. Additionally, the elements of opportunism (Casey & Goldman, 2010) and intelligent opportunism (Liedtka, 1998a) emphasize the need to be actively looking for beneficial change. This is a crucial distinction, as change can be forced upon a strategist, by force or circumstances. Changing course in the face of an obvious need for change does not necessarily indicate flexibility, rather the truly flexible strategic thinker is opportunistic, always aware of where potential changes may prove beneficial.

Flexibility also underlies action learning as an inflexible orientation might cause one to hesitate to look fairly at the results of an action, for fear that the results might indicate changes are necessary. The specific model elements contributing to the theme of action learning included testing (Casey & Goldman, 2010) and hypothesis-driven thinking (Liedtka, 1998a). As noted earlier, action learning is essentially a process of applying the scientific method. Flexibility and openness to new concepts and new methods is critical to any scientist who wishes to remain current in his or her field. The same applies to an iterative strategic thinker who must test strategies and fairly evaluate them constantly.

Finally, the theme of creativity was directly referenced in several models of strategic thinking (Bonn, 2001; Graetz, 2002; Grier, 2012; Heracleous, 1998; Waters, 2011; Yarger, 2008). The process of creation is inherently incompatible with inflexibility. New ideas and associations cannot be formed without some degree of flexibility. Moreover, the willingness to embrace the changes to one's understanding and processes that comes from implementing creative ideas requires an even greater level of flexibility.

As with any characteristic, there is a threshold beyond which flexibility is problematic. The execution of strategy requires long-term consistency, so long as the strategy is still the right approach. The key is to think of this characteristic as *flex*ibility, rather than *break*ability. Put another way: bend, but don't break from a long-term strategy, unless adhering to that strategy is only justified by escalated commitment. *Bending* a strategy allows for adaptation and adjustment of a strategy and its implementation in the face of changing conditions, without abandoning the long-term effort and developing another strategy from scratch.

Inclusiveness. The next mindset characteristic is intellectual inclusiveness, referring to the welcoming of information and opinion from a broad range of sources. The sources of information here could refer to individuals, groups, disciplines of study, or other relevant sources of information and perspective. Inclusiveness is conceptually similar to the openness to experience element of the Five Factor Model of personality (Costa & McCrae, 1992). A mindset favoring exclusion could result in one being hesitant to consider new or unusual sources of information or fearful of having too many voices involved in a discussion. Conversely, maintaining an inclusive mindset allows one to value the holistic understanding that can come from examining an issue from many directions. As with flexibility, inclusiveness could lead to a point of diminishing returns, at which point information overload might lead to a feeling of analysis paralysis (Sugerman, Scullard, & Wilhelm, 2011). However, an inclusive mindset values the potential benefit of a broad perspective, while relying on other processes and judgment to eliminate or ignore information that does not contribute to understanding.

Inclusiveness primarily underlies the themes of (a) breadth and inclusion and (b) systems and synthesis. Specific model elements generating the theme of breadth and inclusion included engaging with diverse perspectives (Yorks & Nicolaides, 2012), assessing trends in divergent domains (Yorks & Nicolaides, 2012), openness to discourse (Waters, 2011), scanning (Casey & Goldman, 2010; McCauley, 2012; Waters, 2011), and openness to many disciplines (Salmoni et al., 2010). Engaging in a broad and inclusive information search requires a welcoming of this kind of search.

Inclusiveness is also critical to systems thinking and synthesis, which differs from traditional analytical processes in that system factors are gathered together for holistic understanding, rather than segmented into divisions handled separately. Specific model elements for this theme were quite tightly clustered around the idea of systems thinking (Casey & Goldman, 2010; Pisapia et al., 2005; Waters, 2011; Yarger, 2008;) and synthesis (Graetz, 2002; Heracleous, 1998; Waters, 2011). Without an inclusive mindset, attempts to synthesize new understanding will clearly be hindered and likely result in nothing truly new. Likewise, one's view of a problem's context (i.e., the system) will be limited to the range of perspective and information already considered relevant.

The need for selecting strategic thinkers with an inclusive mindset was supported not only by analysis of the existing models of strategic thinking, but also by a foundational theory underlying organizational behavior: the behavioral theory of the firm, as presented by Cyert and March (1963). This theory introduced the concept of bounded rationality in the decision making of organizations. Bounded rationality means that managers are faced with information search and processing limits and will not always make decisions that maximize profit perfectly. Instead, managers will rely on closed search processes and decision-making heuristics or established rules that allow them to "satisfice" or reach an acceptable level of performance. Strategic thinking requires going beyond heuristics and the habits of the past to create new strategies for accomplishing different objectives or new levels of performance.

Humility. The final mindset characteristic is intellectual humility. This characteristic refers to a comfort level with being wrong or having an incomplete understanding. Furthermore, comfort with being wrong must be accompanied by the tendency to check oneself, examining issues as if one's understanding is wrong in some way. The importance of humility in leadership

has been cited as an area in need of greater theory and research (Barling, Christie, & Hoption, 2011). However, Morris, Brotheridge, and Urbanski (2005) laid out a strong review of the concept of humility and its role in leadership, defining humility as "a personal orientation founded on a willingness to see the self accurately and a propensity to put oneself in perspective" and emphasizing that humility "involves neither self-abasement nor overly positive self-regard" (p. 1331).

Humility enables objectivity about the self, which is clearly necessary for many of the competencies associated with strategic thinking. Specifically, humility is a necessary pre-condition to the self-criticism and objectivity required for self-awareness/self-control and criticality. The literature suggests that strategic thinkers must have self-awareness of biases and assumptions (Waters, 2011), self-control (Grier, 2012), meta-cognition (Salmoni et al. 2010), questioning (Casey & Goldman, 2010), and reflection (Pisapia et al., 2005). Humility is a crucial component to these competencies, at least as they relate to criticality and awareness of oneself and any group with which one identifies (e.g., specific functional area, organization, industry).

Humility alone does not ensure objective self-awareness, self-control, or criticality, but it is a necessary pre-condition to accepting that one may be biased or clinging to faulty assumptions. Without humility, any change in understanding or alternative viewpoint poses a threat to self-esteem, opening the door to defensiveness and bias in favor of maintaining old beliefs or assumptions. Once open to and comfortable with the position of reasonable self-doubt, behaviors and positions can be examined and, with effort, improved.

Other examples of mindsets. The separation of a mindset (or orientation) from its actualized capability, either behaviorally or cognitively, can be found in other scientific literature. Lumpkin and Dess (1996) conceptualized a variable called entrepreneurial orientation, noting that it is distinct from entrepreneurship itself (that being the act of starting a new business or business venture) and has to do with individual characteristics that predict acting entrepreneurially. Much progress has been made in understanding motivation through research on goal orientation (Payne, Youngcourt, & Beaubien, 2007). In this literature, the act of setting effective goals is differentiated from one's orientation toward what kinds of goals to set (e.g., learning vs. performance goals). Dweck (2006) described a body of research on mindsets related to the self, specifically on the malleability of one's talents and abilities. Her findings suggest that much of success has to do with interpreting a challenge as an opportunity to develop, rather than a threat to reveal one's limitations. Finally, Story and Barbuto (2011) examined the concept of global mindset, describing it as a combination of cultural intelligence and global orientation.

Research on the strategic thinking mindset is in a very nascent stage and is in great need of further refinement, from a definitional, developmental, and measurement perspective. There are some models of strategic thinking that reference the importance of a person's mindset, independent of their thinking competencies. The Pisapia et al. (2005) model, described above, is framed using the term mindset. The Waters (2011) framework uses the term foundation (i.e., something beyond cognitive skills and intelligence), which he describes as having self-awareness of one's own biases and assumptions, including the influence of one's culture, being considerate of ethical and value-related issues, and having an openness to discourse and reflection. Finally, Yorks and Nicolaides (2012) focus on differentiating the idea of a mindset for strategic thinking, apart from the ability, but do not discuss the exact nature of the mindset in much detail.

Yorks and Nicolaides (2012) reference a useful theoretical framework to illustrate how a strategic thinking mindset might differ from other mindsets. The framework was developed from the theory of developmental action inquiry (Torbert, 2004). The Torbert theory details a progression of "action logics" that guide how a leader generally thinks and acts at a series of stages of organizational maturity. Each stage is represented by its own guiding principle. The progression begins with an "opportunist" mindset that is guided by a principle of self-interest and winning as the only concern. The penultimate stage is the "strategist" mindset, in which the focus is on linking valued theories and principles with action in dynamic systems. The Torbert theory presents these mindsets as developmental stages that are passed through over time as a manager matures. Torbert's theory was not developed with specific reference to strategic thinking. Rather, Torbert's theory is an attempt to depict how the mindset of a leader evolves over time, adopting new guiding principles. Therefore, Torbert's theory supports the current research by the notion that there are managerial mindsets that vary and can be developed over time.

In the military domain, Yarger (2008) focused on the importance of understanding the differences between operational planning, national strategy, and national policy, noting that there are subtle, but important, distinctions and that "each also has a different mindset" (p. 8). However, Yarger's explanation of the strategic thinking mindset consisted of the five thinking competencies described earlier. The strategic thinking mindset conceptualized in this work refers to something antecedent to those elements. The strategic thinking mindset reflects a construct in line with the capacity element in Yorks and Nicolaides (2012), while Yarger's work reflects more of the competency element.

Strategic thinking mindset change. Separating the concepts of strategic thinking ability and strategic thinking mindset raises the question of the malleability of strategic thinking mindset; that is, whether the variable is more state-like (i.e., changing over time, transient) or trait-like (i.e., resistant to change, stable). The concept of a strategic thinking mindset being developed is more trait-like than state-like, but still malleable over time. The general domain knowledge concept described by Lievens and Motowidlo (2016) as resulting from "fundamental socialization processes (parenting, schooling, etc.)" (p. 8) reflects how the mindset might also develop. Lievens and Motowidlo also discuss the ways in which the implicit trait policies people hold—the degree to which people perceive traits to be effective in a social situation (see Motowidlo, Hooper, & Jackson, 2006)—result in general domain knowledge when a job's requirements align with those traits. This is similar to the way in which the mindset might be considered a type of trait-based knowledge of the job of a strategic thinker.

The Torbert (2004) theory of developmental action inquiry described above depicts a development sequence of a leader's maturing mindset. A strategic thinking mindset is something that could be similarly developed over time, as one matures and incorporates experiences into a certain worldview. There are other examples of attributes that function between the extremes of state-like (transient) variables and trait-like (stable, unchanging) variables. For example, research on psychological capital (self-efficacy, hope, optimism, and resilience) has demonstrated developmental capacity, while still being relatively stable over short periods of time (Luthans, 2002). However, there may be individual differences in how early or naturally this mindset is adopted. For example, individuals with little openness to experience (Costa &

McCrae, 1992) or tolerance for ambiguity (Budner, 1962) may naturally be inclined to the cognitive requirements of strategic thinking, as described above.

A great deal of literature argues that strategic thinking skills must be developed through participation in the generation of strategy (Casey & Goldman, 2010; Mintzberg, 1994a). This research draws on experiential learning theory (Kolb, 1984), which describes the importance of experiencing a process in order to learn it in the context of adult education. While the thinking competencies for strategic thinking may not develop to maturity until one has a wealth of experience and domain-relevant knowledge, the *mindset* for strategic thinking may be present without the developed skills. In this case, assessing a thinker's mindset may be a key addition to assessments of cognitive ability and other individual difference variables in the prediction of future strategic thinking ability.

The Strategic Thinking Mindset Test

Situational judgment test format. To become a successful strategic thinker, having a mindset for strategic thinking is crucial, in combination with cognitive ability and extensive domain knowledge. The ability to assess that mindset has clear practical and theoretical benefits for academics and practitioners concerned with strategic thinking as a competency. This research seeks to access strategic thinking mindset more directly by developing the Strategic Thinking Mindset Test (STMT) using a situational judgment test (SJT) approach.

Using an SJT format, respondents are assessed through their choices regarding specified problem scenarios. In the STMT, the focus is on a respondent's tendency to adopt a strategic thinking mindset in response to those problems. This tendency is inferred from the courses of action they endorse and reject. The respondent's mindset is accessed through their response to concrete scenarios, rather than relying on their own generic self-assessment. This approach also supports development by providing content upon which to base a counseling session discussion. SJTs are typically produced using specific work-related problem scenarios that represent a domain of performance. To balance applicability to all Army leaders while still rooting scenarios in a specified work domain, the problem scenarios forming the basis of the item content were collected from Company-grade Army officers with no restriction to a particular functional specialty. This enables the developed instrument to be used widely throughout military organizations.

A practical advantage for using an SJT format for the STMT includes the opportunity to assess a wide candidate pool in search of future strategic thinkers in an organization. Higher-fidelity simulations and exercises have their own advantages for assessing actual strategic thinking. However, the costs of trained assessors and complex simulation environments require a restriction of who can be feasibly assessed. Paper-and-pencil SJTs are simpler and cheaper to administer; therefore, SJTs could be used to assess more personnel than would likely be feasible with assessment centers or simulation exercises. Furthermore, assessing whether developing leaders have a strategic thinking mindset, apart from the actualized capability, permits earlier identification of potential, prior to a point when leaders will have had the opportunity to fully demonstrate strategic thinking ability. With earlier identification of potential, mentorship and development can occur for a longer period prior to placement in a strategic-level position. From the test takers' perspective, they can learn from the results of the SJT about the ways in which

their thought processes deviate from what theory says about dealing with strategic problems, allowing them to seek out development on their own and raise their metacognitive awareness, a key element of strategic thinking.

The development and testing of the STMT items occurred in four main stages: Stage 1–Key Incidents, Stage 2–Feedback and Response Options, Stage 3–Response Option Scoring, and Stage 4–Pilot Testing and Construct Validation. In Stage 1, critical incidents were gathered and selected for relevance to the content area of the instrument (i.e., complex, ambiguous problem scenarios at the Army Company-grade officer level), consistent with the critical incident technique (Flanagan, 1954). The remainder of this report will refer to these incidents as "key" incidents, for clarity with a military audience. In Stage 2, Army personnel served as subject matter experts (SMEs) providing feedback on the quality of the scenarios (e.g., fidelity, detail, complexity, and ambiguity), determining which of the mindset characteristics (flexibility, humility, or inclusiveness) were most relevant for each scenario, and generating a range of possible response options. In Stage 3, a new sample of SMEs rated a set of candidate response options for each scenario on effectiveness and level of mindset expression. Finally, in Stage 4, two final groups of Army personnel completed the instrument and a set of additional measures for construct validity evidence.

Research questions and hypotheses. The current research seeks to address two broad research questions. First, given the need for, and potential benefit of, an assessment of a young military officer's mindset for strategic thinking, can an SJT be designed to assess this mindset through *flexibility*, *inclusiveness*, and *humility* characteristics? Secondly, will this test exhibit evidence of construct validity and measurement reliability?

The first research question relates to the structural characteristics of the test. The instrument was developed to measure three distinct mindset characteristics. These characteristics are somewhat interrelated and will likely correlate with each other. The following hypotheses reflect the anticipated nature of the test scores and their relationships with other variables:

> *Hypothesis 1*: An oblique, three-factor model will result in a better fit than competing orthogonal and single-factor models.

Construct-related validity is demonstrated by triangulating the STMT score for each mindset characteristic with measures of conceptually-related constructs (convergent validity) and conceptually unrelated constructs (discriminant validity). Hypotheses 2a–c summarize the evidence regarding convergent validity:

> *Hypothesis 2a*: There will be an inverse relationship between strategic thinking mindset flexibility and resistance to change.
>
> *Hypothesis 2b*: There will be a positive relationship between strategic thinking mindset inclusiveness and work-related openness.
>
> *Hypothesis 2c*: There will be a positive relationship between strategic thinking mindset humility and dispositional humility.

It should be noted that, in the case of Hypothesis 2a, the evidence of convergence is demonstrated by an inverse (negative) relationship, due to the framing of the resistance to change variable. Discriminant validity evidence will be gathered from the relationship between each mindset characteristic and a measure of cognitive ability. Meta-analytic estimates of the relationship between SJT scores and scores from cognitive ability tests showed a mean correlation of .29 (corrected .32; McDaniel et al., 2007). Conceptually, strategic thinking mindset should be less related to cognitive ability than a traditional SJT construct that is closer to true judgment or job knowledge. Therefore, discriminant validity will be inferred from a small or non-significant relationship, given adequate measurement reliability (coefficient alpha). Given the problems associated with hypothesizing the null (Cortina & Folger, 1998), no formal hypotheses are associated with this test of validity.

Test Development

Stage 1 Method

Sample. A sample of 125 Soldiers participated in developing the key incidents, from eight U.S. Army installations across the country. The sample consisted of 104 males and 21 females, very closely matching the gender breakdown across the Active Army (17% female, Maxfield, 2015). The vast majority were Army Captains (85.6%), and represented a wide range of Army functional branches, with no more than 18% of the sample coming from a single branch. A large majority (82.4%) had deployment experience. The most recent deployment was predominately to Afghanistan (45.6%) or Iraq (29.6%). Additional detail on the sample, including the branch representation, is provided in Table 3.

Table 3

Stage 1 Sample Demographics

Gender	Male *n* = 104	Female *n* = 21		
Branch	Intelligence *n* = 21	Engineer *n* = 15	Armor *n* = 12	Infantry *n* = 12
	Field Artillery *n* = 10	CBRN* *n* = 8	Military Police *n* = 7	Logistics *n* = 7
	Aviation *n* = 6	Air Defense *n* = 4	Marines* *n* = 4	Adjutant General *n* = 3
	Medical Services *n* = 3	Health Services *n* = 3	Quartermaster *n* = 3	Signal *n* = 3
	Ordnance *n* = 2	Public Affairs *n* = 1	JAG* *n* = 1	
Rank	Major *n* = 5	Captain *n* = 107	First Lieutenant *n* = 8	Gunnery Sergeant (Marine) *n* = 1
	Staff Sergeant (Marine) *n* = 1	Sergeant (Marine) *n* = 2	Specialist *n* = 1	
Avg. time in rank (months)	Major 53	Captain 27	First Lieutenant 13	Gunnery Sergeant (Marine) 48
	Staff Sergeant (Marine) 36	Sergeant (Marine) 39	Specialist 48	
Most recent deployment	Afghanistan *n* = 57	Iraq *n* = 37	Other nation *n* = 9	Never deployed or none indicated *n* = 22

Note. CBRN = Chemical, Biological, Radiological, and Nuclear; JAG = Judge Advocate General.

Materials and procedure. Participants received three documents: a demographic questionnaire, an instructions packet, and a worksheet. The demographics sheet was completed and returned independently. Participants were provided an instructions packet (see Appendix B) that provided detailed explanations, tips, and examples for writing the key incident(s). The participants were briefed on the content of this packet and were given time to review it while they thought of their incident(s). Participants were free to ask clarifying questions throughout the session.

The participants used the worksheet to write their key incident(s). Participants were asked to provide at least one incident. However, if time allowed, the worksheet was designed to accommodate two incidents. In most cases, participants only provided one incident. Participants reported their incidents in three sections: Situation, Problem-Solving Approach, and Outcome. The situation section formed the core content for the test scenarios, while the approach and outcome sections provided context with which to develop the response options. Participants were instructed to speak in third-person narrative using false names of people and specific places (e.g., a particular base) and only broad time frames, if necessary.

In addition to writing the incident, participants were asked to provide two additional pieces of information: the type of thinking required by the incident, using Yarger's (2008) model, and a rating of the degree to which the problem had been effectively addressed (1 = *Very Ineffective/Detrimental* to 5 = *Highly Effective/Successful*). This was done to provide further contextual information on how the participant viewed the incident and whether/how to use each incident as a scenario. These ratings were not included for any quantitative analysis purpose.

Stage 1 Results

A total of 144 key incidents were collected, of which 59 were viable for the particular nature of the test. The incident descriptions were converted into open-ended scenarios by removing the Problem-Solving Approach and Outcome sections, leaving only the Situation section, which formed the basis of the scenario. Occasionally, content from the Approach and Outcome sections was brought into the scenario, either because it provided more information about the situation than the problem-solving approach or because it served to extend the complexity or ambiguity of the situation. The author also edited the scenarios as needed for grammar, spelling, clarity, and removal of any potentially identifying information that was not censored by the original writer of the incident.

Two PhD-level research psychologists evaluated each of the 59 scenarios with the following criteria in mind: a clearly stated problem depicted with contextual detail, ambiguity with regard to the correct way to address the problem, and opportunity for a Company-grade officer to make choices varying in flexibility, inclusiveness, and/or humility. After this extensive incident and scenario review process, 32 scenarios were selected for further development as items in the scale.

Stage 2 Method

Sample. A total sample of 75 Soldiers with Company-grade deployment experience from four participating installations was surveyed as subject matter experts (SMEs), for the purposes of evaluating the scenarios and providing realistic response options. The status of the Soldiers as an "expert" sample refers to their familiarity with and knowledge of realistic intellectual and behavioral responses that a Soldier might have to a given scenario. The sample consisted of 69 males, 5 females (with one non-respondent), predominately Captains (74.7%), representing a range of Army functional areas, although a large proportion came from the infantry branch (41.3%). Additional detail on the sample is provided in Table 4.

Table 4

Stage 2 Sample Demographics

Gender	Male *n* = 69	Female *n* = 5		
Branch	Infantry *n* = 31	Armor *n* = 10	Intelligence *n* = 7	Field Artillery *n* = 7
	Logistics *n* = 4	Engineer *n* = 3	Adjutant General *n* = 2	CBRN *n* = 1
	Aviation *n* = 2	Military Police *n* = 1	Signal *n* = 1	JAG *n* = 1
	Simulation Operations *n* = 1	Public Affairs *n* = 1	Army Medical Specialist Corps *n* = 1	Army Nurse Corps *n* = 1
Rank	Lieutenant Colonel *n* = 1	Major *n* = 15	Captain *n* = 56	First Lieutenant *n* = 2

Note. CBRN = Chemical, Biological, Radiological, and Nuclear; JAG = Judge Advocate General.

Note that not all participants reviewed each of the scenarios. A significant amount of time was required, per scenario, to read, consider, and provide feedback and response options. A unique set of scenarios was provided to participants in each data collection session. Due to varying participation rates, and the relative need (or lack thereof) for additional response options

on certain scenarios, the number of participants sampled for each scenario ranged from 17 to 25 (23% to 33% of the total sample).

Materials and procedure. Participants were given three documents: a brief demographic questionnaire, a construct definition handout, and a scenario review and feedback packet. The demographics sheet was completed and returned independently. The participants then reviewed and listened to the researcher explain the nature of the research and the meaning of the three mindset characteristics. Participants were instructed to familiarize themselves with the definitions of flexibility, humility, and inclusiveness as indicated on the definitions sheet (see Appendix C). The researcher reviewed the content of the definitions sheet with participants and elaborated on the meaning and derivation of each characteristic. Participants were free to ask clarifying questions throughout the session. Finally, participants reviewed their assigned scenarios in the *Scenario Review and Feedback Packet* (see Appendix C). Each packet consisted of approximately 10 scenarios for review. The exact number, group, and ordering of the scenarios changed for each session, to avoid order effects as well as to ensure sufficient feedback was gathered for each scenario.

The researcher explained to the participants the need for their contextual familiarity with military procedures and problem solving to evaluate the scenarios. Participants were instructed to write whatever feedback occurred to them, with particular emphasis on the degree to which the scenario reflected a sufficiently ambiguous and complex problem scenario, the realism of any details provided (e.g., a newly-promoted Captain being placed in a certain type of position), and the need for any additional clarifying details (although in some cases, the lack of such details is the root of the scenario's ambiguity).

Participants were also asked to consider each of the three strategic thinking mindset characteristics (flexibility, humility, and inclusiveness) and select the one that was the most relevant to the scenario. That choice represented a vote for the most appropriate characteristic to assess with that scenario. After providing the scenario feedback, participants then provided up to ten feasible response options, with instructions to vary their options by the mindset characteristics (i.e., providing a highly flexible response, a highly inflexible response, and a neutrally flexible response). It was emphasized that they need not restrain themselves to responses that would fully address the problem or that they would even consider to be good responses. Rather, they were instructed to describe as many *feasible* responses as they could (i.e., responses that someone in the Army might realistically make). Finally, they were instructed that their responses might and ought to include cognitive responses, or ways in which the person depicted in the scenario might think about the scenario, as well as act.

The initial data collection session for this stage was conducted in a small group discussion format. After reviewing the scenarios, the researcher facilitated a group discussion among all participants about each scenario for the final 15 minutes of the session. This approach was abandoned for the remaining data collections. The primary reason was the need by many participants to use the available time to get through all the scenarios. Furthermore, the group discussions tended to result in very little new insight through cross-talk.

Stage 2 Results

The first step in analyzing the data for this stage was to tally the mindset characteristic votes for each scenario. In several cases, respondents selected two characteristics for a scenario. In this case, both were counted as a vote. The winning characteristic was documented and, in cases where there was a tie or close vote, a back-up characteristic was also noted. Initially, the voting process resulted in 17 scenarios for measuring flexibility, 5 scenarios for measuring humility, and 9 scenarios for measuring inclusiveness. One of the 32 scenarios, having to do with the threat of improvised explosive devices (IEDs) on a convoy route, was eliminated due to feedback that it was not only insufficiently ambiguous, but also in consideration of the stress-inducing potential of such an item.

The goal of this process was to have an even number of scenarios (approximately 10) measuring each of the three mindset characteristics. Therefore, the process of assigning the characteristic for each scenario proceeded iteratively. First, all scenarios with a tied (or nearly tied) vote that included flexibility were assigned to the other characteristic. Likewise, the 10 scenarios that received the most votes for flexibility retained that assignment. The remainder were assigned according to the number of votes between the two remaining characteristics. When necessary, scenarios were altered to more closely match them to the assigned characteristic. The culmination of this process was 10 scenarios for flexibility, 10 for humility, and 11 for inclusiveness.

The next step was to review, edit, and assign the response options for each scenario. The generation of these response options relied heavily on the content provided by the SMEs in Stage 2, particularly when it was indicated that a response option was meant to reflect a high or low level of a certain characteristic. There were many response options provided by the participants from Stage 2. The number of responses ranged from 24 to 74, with an average of 40 per scenario. Although they were not all appropriate or usable for various reasons (e.g., options that were illegal, immoral, clearly against policy, beyond an individual's realistic control, or otherwise obviously bad responses), the range of responses provided a clear sense of the types of responses Army personnel would find realistic. Unsurprisingly, many of the response options were similar, giving a further sense of what "common sense" among Army officers might dictate.

In selecting and producing a quality set of response options for rating in Stage 3, several factors were considered. The first consideration was whether a response option clearly indicated a high or low level of the characteristic assigned to that scenario, without being explicit (e.g., "remain flexible when talking with the host national"). A second consideration was whether the response option represented a reasonable, but not certain, chance of resolving the problem. A third consideration was whether the option was a realistic/feasible way for someone to respond. A final consideration was whether the response option matched the level of specificity of other options for that scenario, regarding scope of action and time. All these factors were considered in selecting response options for each scenario. The author also used the provided response options as a base from which to judge the appropriateness of newly written or heavily-revised responses that would fill in the necessary number and type of responses needed for Stage 3. On average, no more than one option per scenario needed to be newly composed or heavily revised

by the author. In certain cases, scenario details were tweaked to allow for a wider range of possible responses.

For each scenario, six (in a few cases, seven) response options were selected or written to indicate a positive, negative, or neutral expression of the characteristic (two response options for each level of expression). The neutral and negative response options relied on their juxtaposition with the positive response option. A different approach to the negative response options might have been to select response options that are more transparently indicative of an opposing characteristic (e.g., flexibility vs. rigidity, inclusiveness vs. isolation or exclusion, and humility vs. arrogance or condescension). This approach, however, would have risked making the items too transparent and restrict variance in the test scores. Although they might be endorsed by some (e.g., those who would proudly reject the value of flexibility, inclusiveness, or humility), these response options would most likely exacerbate the restriction of range in the item score distribution.

This process occurred through multiple iterations of revision between the author and other PhD research psychologists employed by the Army who brought insight into Army leadership doctrine, training, and practice. This marked the end of Stage 2, with a collection of 31 scenarios, each accompanied now by a set of six response options to be evaluated for effectiveness and level of the assigned characteristic in Stage 3.

Stage 3 Method

Sample. A total sample of 224 Army personnel (primarily 1LTs and CPTs, some Non-Commissioned Officers) with deployment experience were sampled as SMEs for this stage (see Table 5). As with the previous stage, the participants' status as expert is in relation to their understanding of Army Soldier behavior and problem solving in context. As with Stage 2, each participant evaluated a subset of the total scenario pool, in this case about half. The sample size for each scenario ranged in size from 68–90 (M = 83.8). Also, a group of four research psychologists (three PhD-level, one Master's-level) familiar with Army leadership doctrine and scale development methodology evaluated the scenarios to supplement the Army SME ratings with an alternative perspective.

Table 5

Stage 3 Sample Demographics

Gender	Male *n* = 203	Female *n* = 20		
Branch	Field Artillery *n* = 48	Air Defense Artillery *n* = 38	Intelligence *n* = 46	CBRN *n* = 24
	Military Police *n* = 20	Infantry *n* = 17	Armor *n* = 5	Engineer *n* = 5
	Aviation *n* = 4	Signal *n* = 4	Cavalry *n* = 3	Medical *n* = 2
	EW *n* = 2	Ordnance *n* = 1	Quartermaster *n* = 1	Transportation *n* = 1
	Adjutant General *n* = 1			
Rank	Corporal *n* = 1	Sergeant *n* = 3	Staff Sergeant *n* = 44	Sergeant First Class *n* = 5
	Warrant Officer *n* = 1	2nd Lieutenant *n* = 4	1st Lieutenant *n* = 49	Captain *n* = 114
	Major *n* = 2			

Note. CBRN = Chemical, Biological, Radiological, and Nuclear; EW = Electronic Warfare. One participant did not report their branch.

Materials and procedure. Participants were provided with a similar set of materials as in Stage 2. The demographics questionnaire and definitions sheet were the same as in the previous stages. Participants also received a rating packet consisting of approximately 15 scenarios, each of which was followed by the list of response options twice (for an example, see Appendix D). The first time through the response options, participants used a 5-point Likert-type scale to indicate how strongly they agreed or disagreed that a response option would be an effective way to address the problem (1 = *Strongly Disagree* to 5 = *Strongly Agree*). Afterward, the participants rated each response option a second time, using the same scale, this time

indicating how strongly they agreed or disagreed that the response would reflect a mindset of humility/inclusiveness/flexibility, as noted for that scenario.

Participants were briefed on the nature and objectives of the project in a fashion similar to Stage 2. Additional emphasis was placed on their understanding the difference between rating the effectiveness of a response option (i.e., how well would it address the problem scenario presented?) and how much it reflected the mindset characteristic (i.e., regardless of whether it is a good response, is it a flexible/humble/inclusive response?). Notably, the expertise of the SME sample is considerably stronger in regard to the likely effectiveness of a response option as compared to the level of characteristic shown. Therefore, considerable time was spent in the briefing of the research about the meaning of flexibility, humility, and inclusiveness in this study. Similar to Stage 2, the order of presentation in the scenario packet varied. For Stage 3, the order was carefully counter-balanced to avoid the order effects related to survey fatigue or contamination from a previous scenario.

Stage 3 Results

Each scenario and its accompanying response options were reviewed with a goal to identify the best scenarios (6–8 for each characteristic) to use in pilot testing. Recall that the result of Stage 2 was a set of six response options for each scenario, two of which were written to represent a positive expression of the mindset characteristic, two representing a negative expression, and two representing a neutral expression.

For a scenario to become a pilot-test item, the scenario needed a set of four response options with similar effectiveness ratings, one option to be positively keyed, one option to be negatively keyed, and two to be unkeyed, neutral options. This required a careful examination of the ratings for each response option (186 in total), as well as an evaluation of the scenarios with the best set of options. On an individual scenario basis, two exclusion criteria were applied: (a) a correlation $\geq$ .95 between ratings of effectiveness and the mindset characteristic (< 1% of ratings removed), and (b) zero variance in ratings of the mindset characteristic (3.3% of ratings removed). Extremely high correlations between effectiveness and mindset indicate that the participants did not differentiate between the elements they were being asked to rate. Similarly, a participant who did not vary his or her ratings of flexibility, for example, for all response options to a scenario contributed no value to the differentiation of the responses.

The primary criteria for evaluating a response option was the mean characteristic expression ratings given by the SMEs and the psychologists. Additionally, the standard deviation of the expression ratings was considered, with smaller deviations indicating greater rater agreement. Given the overall positive skew to the characteristic ratings (M = 3.52, SD = 1.3; 5-point scale), a rule of thumb was adopted for evaluating whether each option matched its intended level of characteristic expression. For positive expression options, a mean characteristic rating greater than 4.25 was a good match. Likewise, for negative expression options, a mean characteristic rating below 2.5 was a good match. For neutral options, mean characteristic ratings between 3.25 and 3.75 were considered good matches. For each scenario, the set of options had to be considered as well. Scenarios with a larger range between the mean ratings for the positive and negative expression options were favored. Likewise, scenarios with options scoring fairly equally on effectiveness were favored. Discrepancies between the ratings

from Army SMEs and the psychologists were examined to further explore the level of agreement about how each option expressed the characteristic.

The process of selecting the best four response options for each scenario, and in turn selecting the best scenarios for each characteristic, proceeded according to these criteria. When multiple options could work for a particular keying, preference was given to the one with a smaller standard deviation and less discrepancy between the SME and psychologist samples. In cases where ratings for an intended keyed option did not support the intention, but a neutral response option aligned better, it was substituted. Ultimately, seven scenarios were chosen as items to measure inclusiveness, seven scenarios as items to measure flexibility, and seven scenarios as items to measure humility. In Tables 6–8, the data for each response scenario and its response options are presented, along with which scenarios and options were selected and how the options were keyed.

Table 6

Response Option Keying Data–Flexibility Items

			Flexibility				Effectiveness			
			SMEs		Psychologists		SMEs			
S	Options		M	*SD*	M	*SD*	M	*SD*	Key	Pilot #
1	1	Negative	3.41	1.2	2.25	0.8	3.63	1.2	Negative	P20
	2	Negative	4.46	0.8	4.00	1.2	4.60	0.6		
	3	Neutral	3.32	1.2	4.00	0.0	3.36	1.2		
	6	Neutral	3.21	1.3	2.50	1.1	3.49	1.1	Neutral	
	4	Positive	2.88	1.4	3.50	1.7	3.22	1.3	*Neutral*	
	5	Positive	4.14	0.9	4.25	0.4	4.12	0.9	Positive	
2	1	Negative	3.59	1.2	3.00	0.7	3.53	1.1		P8
	5	Negative	2.82	1.2	2.50	0.5	2.94	1.1	Negative	
	2	Neutral	3.05	1.2	1.25	0.4	4.07	0.9	Neutral	
	4	Neutral	3.85	1.1	4.50	0.5	3.49	1.2	Neutral	
	3	Positive	4.30	0.7	4.50	0.5	3.93	1.0	Positive	
	6	Positive	4.40	0.9	5.00	0.0	4.48	0.7		
14	1	Negative	3.82	1.3	4.00	1.2	4.00	1.1		P9
	4	Negative	2.70	1.2	2.00	0.7	3.23	1.3	Negative	
	5	Neutral	3.02	1.1	3.25	0.8	2.91	1.2	Neutral	
	3	Neutral	4.36	0.8	3.75	0.4	4.22	0.9	*Positive*	
	2	Positive	3.85	1.2	3.75	0.4	3.74	1.1		
	6	Positive	3.56	1.3	4.00	0.7	3.01	1.2	*Neutral*	
10	1	Negative	3.51	1.2	3.75	0.4	3.20	1.4		P2
	5	Negative	2.44	1.3	3.00	1.4	2.64	1.4	Negative	
	3	Neutral	3.83	1.0	2.00	0.7	3.95	1.1	Neutral	
	4	Neutral	3.58	1.2	2.25	0.8	3.57	1.1	Neutral	
	2	Positive	3.96	0.9	4.25	0.4	3.85	1.0	Positive	
	6	Positive	3.86	1.2	3.75	0.8	3.48	1.3		
19	4	Negative	4.04	1.0	4.00	0.7	3.74	1.3		P13
	2	Negative	2.48	1.0	2.75	0.4	3.21	1.0	Negative	
	5	Neutral	2.93	1.1	3.25	0.8	3.06	1.2	Neutral	

			Flexibility				Effectiveness			
			SMEs		Psychologists		SMEs			
S	Options		M	*SD*	M	*SD*	M	*SD*	Key	Pilot #
	6	Neutral	2.81	1.2	2.75	0.8	2.58	1.3	Neutral	
	3	Positive	2.67	1.2	2.75	1.1	2.94	1.1		
	1	Positive	3.89	0.9	3.75	0.4	4.03	0.7	Positive	
23	3	Negative	2.29	1.1	1.75	0.8	2.98	1.2		P16
	4	Negative	3.01	1.3	3.00	0.7	3.56	1.1	*Neutral*	
	5	Neutral	2.10	1.1	1.50	0.5	2.53	1.2	*Negative*	
	6	Neutral	4.16	0.9	4.25	0.4	4.06	1.0		
	1	Positive	3.58	1.2	3.75	0.8	3.16	1.3	*Neutral*	
	2	Positive	4.18	0.8	4.25	0.4	4.24	0.9	Positive	
26	2	Negative	2.13	1.2	1.25	0.4	2.60	1.2	Negative	P5
	6	Negative	3.14	1.2	2.50	0.5	3.02	1.2		
	4	Neutral	3.69	1.0	3.00	1.0	3.97	0.9	Neutral	
	5	Neutral	2.88	1.2	2.25	1.1	3.28	1.2	Neutral	
	1	Positive	3.65	1.3	4.50	0.5	3.16	1.1		
	3	Positive	3.88	0.8	4.00	0.7	3.77	0.8	Positive	
29	1	Negative	2.66	1.3	2.25	1.1	3.93	1.0	Negative	P21
	3	Negative	3.96	1.1	4.25	0.4	4.09	1.1		
	2	Neutral	4.32	0.7	4.50	0.5	4.32	0.9		
	6	Neutral	3.25	1.3	2.50	0.9	3.52	1.1	Neutral	
	4	Positive	4.46	0.8	4.50	0.5	4.36	0.8	Positive	
	5	Positive	3.97	1.0	4.00	0.7	4.28	0.8	*Neutral*	

Note: Pilot # refers to the re-numbering of items for the Stage 4 pilot test. Empty cells under Key indicate response options that were not retained for that item. Italics indicate an option that was changed from its original designation based on the Stage 3 data.

Table 7

Stage 3 Response Option Keying Data–Humility Items

			Humility				Effectiveness			
			SMEs		Psychologists		SMEs			
S		Options	M	*SD*	M	*SD*	M	*SD*	Key	Pilot #
3	1	Negative	2.08	1.3	2.00	1.2	2.12	1.2	Negative	P7
	6	Negative	3.54	1.1	3.25	0.8	4.07	0.9		
	3	Neutral	3.84	1.0	3.75	0.4	4.00	1.0	Neutral	
	5	Neutral	3.39	1.2	3.25	1.1	4.10	0.9	Neutral	
	4	Positive	4.37	0.9	4.25	0.8	4.39	0.8	Positive	
	2	Positive	4.31	1.0	3.50	1.5	4.35	0.7		
	7	Positive	4.30	0.9	4.25	0.4	4.25	1.0		
13	2	Negative	3.33	1.3	2.75	1.8	3.73	1.1	*Neutral*	P10
	5	Negative	3.20	1.1	2.25	0.4	3.92	1.2	Negative	
	3	Neutral	3.59	1.1	3.00	0.7	4.00	1.1		
	1	Neutral	3.73	1.1	2.75	0.8	3.61	1.2	Neutral	
	4	Positive	4.42	0.9	5.00	0.0	4.10	1.0	Positive	
	6	Positive	3.56	1.2	5.00	0.0	2.48	1.2		

17	2	Negative	2.94	1.26	2.25	0.43	3.90	0.98	*Neutral*	P3
	6	Negative	2.14	1.18	1.75	0.43	2.53	1.30	Negative	
	3	Neutral	4.01	1.01	4.25	0.83	4.27	0.92		
	5	Neutral	3.78	1.08	3.25	0.83	3.55	1.19		
	4	Positive	3.65	1.08	4.75	0.43	3.07	1.32	Positive	
	1	Positive	3.77	1.11	3.00	0.71	4.55	0.72	*Neutral*	
21	2	Negative	2.91	1.2	3.00	1.0	2.78	1.3	Negative	P12
	4	Negative	3.36	1.2	3.50	1.1	3.18	1.4		
	3	Neutral	3.55	1.1	3.50	0.5	3.46	1.3	Neutral	
	5	Neutral	3.49	1.2	3.75	1.1	2.75	1.4	Neutral	
	1	Positive	4.08	1.0	3.75	1.1	4.13	1.1	Positive	
	6	Positive	3.75	1.2	4.50	0.5	3.66	1.2		
24	1	Negative	2.97	1.2	1.50	0.5	4.27	0.7		P4
	3	Negative	2.61	1.1	2.00	0.0	3.11	1.1	Negative	
	4	Neutral	3.43	1.1	3.50	0.5	3.89	1.0	Neutral	
	6	Neutral	3.58	1.0	2.75	0.8	3.86	0.7	Neutral	
	2	Positive	3.97	0.9	4.25	0.4	3.43	0.9	Positive	
	5	Positive	3.68	1.2	4.50	0.5	2.88	1.3		
30	1	Negative	3.53	1.3	2.50	0.9	4.34	0.8	*Neutral*	P19
	2	Negative	2.53	1.2	2.50	0.5	3.15	1.3	Negative	
	4	Neutral	3.43	1.1	2.75	0.8	4.07	0.9	Neutral	
	5	Neutral	2.63	1.3	1.50	0.5	2.46	1.2		
	3	Positive	3.77	1.1	3.50	0.9	3.54	1.1	Positive	
	6	Positive	3.78	1.1	3.25	1.3	4.21	0.9		
32	1	Negative	2.90	1.3	1.75	0.8	3.46	1.1	Negative	P11
	5	Negative	2.93	1.2	2.00	1.0	3.73	1.0		
	2	Neutral	3.17	1.2	2.25	1.1	3.61	1.2	Neutral	
	4	Neutral	2.97	1.2	2.50	1.1	2.69	1.2	Neutral	
	3	Positive	4.20	1.0	4.75	0.4	4.27	0.9	Positive	
	6	Positive	3.93	1.0	3.75	1.1	3.27	1.3		

Note: Pilot # refers to the re-numbering of items for the Stage 4 pilot test. Empty cells under Key indicate response options that were not retained for that item. Italics indicate an option that was changed from its original designation based on the Stage 3 data. Scenario 3 included seven options for scoring.

Table 8

Response Option Keying Data–Inclusiveness Items

			Inclusiveness				Effectiveness			
			SMEs		Psychologists		SMEs			
S	Options		M	*SD*	M	*SD*	M	*SD*	Key	Pilot #
4	1	Negative	2.85	1.2	3.50	1.5	3.35	1.1	Negative	P17
	5	Negative	3.42	1.2	3.00	1.0	3.60	1.1		
	3	Neutral	3.07	1.2	2.75	1.1	3.25	1.2	Neutral	
	6	Neutral	3.43	1.2	2.75	1.3	3.39	1.2	Neutral	
	2	Positive	3.95	0.9	4.00	0.0	4.13	0.8		
	4	Positive	4.41	0.9	4.75	0.4	3.95	1.2	Positive	
6	1	Negative	1.79	1.1	1.00	0.0	2.31	1.3	Negative	P1
	4	Negative	3.25	1.2	2.33	0.9	3.74	1.1	*Neutral*	
	3	Neutral	3.87	1.1	3.50	0.5	4.09	1.0		
	5	Neutral	3.98	1.0	3.50	0.9	3.94	1.0		
	2	Positive	3.16	1.3	3.25	1.1	3.25	1.3	*Neutral*	
	6	Positive	3.97	1.0	3.50	1.1	3.97	1.1	Positive	
8	1	Negative	2.19	1.3	1.50	0.9	2.78	1.3	Negative	P15
	4	Negative	2.99	1.1	2.25	0.8	3.78	1.0		
	3	Neutral	3.66	1.1	3.25	0.8	3.53	1.3	Neutral	
	5	Neutral	3.79	1.1	3.50	0.9	3.44	1.1	Neutral	
	2	Positive	4.36	0.7	4.25	0.4	3.47	1.1	Positive	
	6	Positive	4.27	0.9	4.75	0.4	4.45	0.8		
9	2	Negative	2.24	1.2	1.25	0.4	2.98	1.0	Negative	P18
	5	Negative	3.33	1.1	2.75	0.8	4.30	0.9	*Neutral*	
	1	Neutral	3.69	1.2	3.00	1.2	3.59	1.2	Neutral	
	6	Neutral	3.97	1.1	2.75	1.1	3.92	1.1		
	3	Positive	4.59	0.6	4.75	0.4	4.22	1.0	Positive	
	4	Positive	4.57	0.7	4.00	0.7	4.60	0.7		
11	1	Negative	2.42	1.3	2.00	0.7	3.39	1.3	Negative	P14
	3	Negative	4.07	1.0	3.75	0.4	4.00	1.2		
	2	Neutral	3.40	1.2	2.25	0.8	3.83	1.1	Neutral	
	4	Neutral	2.39	1.2	2.50	0.5	2.61	1.2	Neutral	
	5	Positive	4.57	0.8	4.67	0.5	4.06	1.1	Positive	
	6	Positive	4.49	0.8	5.00	0.0	4.13	1.0		
15	4	Negative	2.83	1.1	2.75	1.1	2.96	1.1	Negative	P6
	5	Negative	3.75	1.1	3.50	0.5	3.55	1.0		
	2	Neutral	3.05	1.4	3.25	0.8	3.52	1.2	Neutral	
	6	Neutral	3.02	1.3	2.50	1.1	2.85	1.2	Neutral	
	1	Positive	4.55	0.8	5.00	0.0	4.40	0.7	Positive	
	3	Positive	4.47	0.8	4.50	0.9	4.32	0.8		

Note: Pilot # refers to the re-numbering of items for the Stage 4 pilot test. Empty cells under Key indicate response options that were not retained for that item. Italics indicate an option that was changed from its original designation based on the Stage 3 data.

Stage 4 Method

Sample. A total of 177 Army personnel participated in this stage to complete the pilot test of the STMT and the construct validation measures. Twenty-six additional participants were removed from the sample due to unscoreable SJT responses. The total sample was split into two groups to prevent survey fatigue. Each group completed the pilot test of the STMT. In addition, Group 1 (n = 123, 84% male) completed measures for testing Hypotheses 2a–c, while Group 2 (n = 54; 81% male) completed supplemental measures for construct validity evidence. Additional demographic information on each group is provided in Table 9.

Table 9

Stage 4 Sample Demographics

Gender	Male	Female		
Group 1	n = 103	n = 20		
Group 2	n = 44	n = 10		
Branch	Infantry	Engineer	Military Police	CBRN
Group 1	n = 11	n = 18	n = 13	n = 7
Group 2	n = 12	n = 12	n = 10	n = 8
	Logistics	Intelligence	Transportation	Armor
Group 1	n = 20	n = 3	n = 13	n = 0
Group 2	n = 0	n = 1	n = 0	n = 4
	Adjutant General	Quartermaster	Medical Services	Field Artillery
Group 1	n = 2	n = 4	n = 1	n = 1
Group 2	n = 0	n = 3	n = 1	n = 1
	Public Affairs	Ordnance	Cavalry	Signal
Group 1	n = 1	n = 8	n = 0	n = 8
Group 2	n = 0	n = 0	n = 1	n = 1
	ADA			
Group 1	n = 13			
Group 2	n = 0			
Rank	2nd Lieutenant	1st Lieutenant	Captain	
Group 1	n = 6	n = 26	n = 91	
Group 2	n = 4	n = 28	n = 22	

Note. CBRN = Chemical, Biological, Radiological, and Nuclear; ADA = Air Defense Artillery.

Materials and procedure. Participants completed the same demographics sheet used in previous stages. All participants in both groups completed the pilot-test version of the STMT. In both groups, participants were instructed to read each scenario and then identify the most and least likely options they, personally, would choose, from the options presented. The STMT has an 11th grade reading level, as indicated by the Flesch-Kincaid Grade Level statistic (11.2), and a Flesch Reading Ease statistic of 50.3 (roughly equivalent to 10th to 12th grade). Although quite high, this reading level is appropriate considering the subject matter of the test and the education level of the test-taking population (in 2014, 90.6% of active duty officers had a high school diploma or higher; Office of the Deputy Assistant Secretary of Defense for Military Community and Family Policy, 2014).

In addition, participants completed five other measures. The additional measures were always completed after the STMT, to prioritize the completion of the STMT and avoid priming effects. To test Hypothesis 2a, that the flexibility subscale of the STMT would be inversely related to resistance to change, participants in Group 1 completed the Resistance to Change Scale (RCS; Oreg, 2003). The RCS is an 18-item self-report measure with items rated using a 6-point Likert-type agreement scale (1 = *Strongly Disagree* to 6 = *Strongly Agree*). An example item is, "Once I've made plans, I'm not likely to change them." The RCS is scored by calculating the mean item rating, after any reverse-scored items are recoded. The observed coefficient alpha for the RCS in the current sample was .84.

To test Hypothesis 2b, that the inclusiveness subscale of the STMT would be positively related to work-related openness, participants in Group 1 completed the Work-Related Openness Scale (WROS; Socin, 2008). The WROS is a 30-item self-report measure with items rated using a 6-point Likert-type agreement scale (1 = *Strongly Disagree* to 6 = *Strongly Agree*). The scale contains five facets: Fantasy, Feelings, Actions, Ideas, and Values, each of which are measured with six items. For example, "I often think of a wide range of possible ways to complete a work-task" is an Ideas item.

To test Hypothesis 2c, that the humility subscale of the STMT would be positively related to dispositional humility, participants in Group 1 completed the Dispositional Humility Scale (DHS; Landrum, 2011). The DHS is a 17-item self-report measure with items rated using a 6-point Likert-type agreement scale (1 = *Strongly Disagree* to 6 = *Strongly Agree*), all beginning with the item stem "In general, I like people who…." For example, "In general I like people who… are willing to take others' advice and suggestions when given." The DHS is scored by calculating the mean item rating, after any reverse-scored items are recoded. The observed coefficient alpha for the DHS in the current sample was .96.

Participants in Group 2 completed the 16-item ICAR Sample Test from the International Cognitive Ability Resource (ICAR) group[1] (Condon & Revelle, 2014). The test measures general cognitive ability with four multiple-choice item types (four items each): Letter and Number Series, Matrix Reasoning, Three-Dimensional Rotation, and Verbal Reasoning. A Letter and Number Series item presents a sequence of five letters or numbers with an underlying

[1] ICAR test items are available as part of the public domain, with some restrictions on use to ensure legitimate scientific purposes. A summary of the current project and how the items would be used was submitted to the group. Access and use of the items was approved.

sequential logic. Respondents must correctly identify the next letter or number in the logical sequence from among six options. Similarly, a Matrix Reasoning item presents a 3x3 grid of shapes with an underlying logic, but one missing cell in the grid. Respondents must correctly identify the missing shape in the grid from among six options. A Three-Dimensional Rotation item presents an image of a six-sided cube. Three of the sides are shown and feature a distinct symbol. Respondents must consider each of six cubes and correctly identify the one cube that is a possible physical rotation of the given cube. A Verbal Reasoning item presents a word problem, the answer to which is one of four options presented. To illustrate, one item presents a series of statements about the relative heights of three boys. The respondent must then consider four statements about the boys' heights and correctly identify which of the four is true. The scale is scored as a percentage of the items answered correctly. The observed alpha for the test was .79.

Finally, participants in Group 2 also completed the Strategic Thinking Questionnaire (STQ; Pisapia et al., 2005). The STQ is a 25-item self-report measure with items rated using a 5-point frequency scale (1 = *Almost Never* to 5 = *Almost Always*) in regard to how one thinks and acts in the context of facing a difficult problem or dilemma. The scale has three components measuring *Reframing* (nine items), *Systems Thinking* (seven items), and *Reflection* (nine items). A sample item for *Reframing* is "I rethink the situation from another point of view." A sample item for *Systems Thinking* is "I define the entire problem before breaking it down into parts." A sample item for *Reflection* is "I consciously look for similarities between my past experiences and the current problem." The STQ is scored separately by component, calculating the mean item rating, after any reverse-scored items are recoded. The observed coefficient alphas for each component in the current sample were: *Reframing* (.53), *Systems Thinking* (.62), and *Reflection* (.60).

Stage 4 Results

Item scoring followed the approach used by Motowidlo et al. (1990) whereby the options chosen for most and least likely (or best/worst) result in integer scores between −2 and 2. The system is straightforward: positive and negative keyed options each give a positive or negative score of 1 or −1, respectively, depending on whether they are endorsed (most likely/best) or rejected (least likely/worst). Neutral options give a score of 0. Each item score is thus a combination of the score for endorsement and rejection.

Standard practice in test construction is to report estimates of measurement reliability. Coefficient alpha is a standard metric for estimating reliability, treating each item as an individual test administration of its own and comparing their consistency to estimate the reliability of the test. In regard to SJTs in particular, internal consistency reliability is considered an inadequate index for reliability. As described by Whetzel and McDaniel (2009), SJT items are typically construct heterogeneous, which interferes with factor loadings and the assessment of scale reliability through homogeneity. Whetzel and McDaniel suggest test-retest reliability or parallel forms approaches instead. The nature of the current sample precluded a test-retest format, and establishing parallel forms was infeasible as well. Therefore, internal consistency estimates are provided, with a clear expectation that they underestimate the reliability of the test, due to the nature of these metrics and the construct heterogeneity of SJT items.

Coefficient alpha reliability estimates for the STMT items were all quite low, by traditional standards of scale construction, when all pilot test items were included: inclusiveness (α = .12), flexibility (α = .19), humility (α = .18), and overall (α = .36). Four items for each characteristic were selected to remain in the item set, on the basis of item-total correlation and item-scale correlation. The refined item set shows improved reliability estimates: inclusiveness (α = .34), flexibility (α = .31), humility items (α = .18), and overall (α = .34).

In Tables 10 and 11, the intercorrelation of scores on all items and the 12-item final scale are provided. Due to the generally low intercorrelations, and the violations of univariate and multivariate normality, the data are not factorable. Therefore, Hypothesis 1 could not be validly tested. For the remaining analyses, scale scores, alphas, and intercorrelations will use the 12-item form of the test. In Tables 12–14, the validity correlations are shown at the item level and scale level using the final 12-item scale.

Table 10 *Interitem Correlation Matrix (All Items)*

			1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24
1		Item 1																								
2	Inclusiveness	Item 6	.08																							
3		Item 11	-.04	.01																						
4		Item 14	.10	.07	-.12																					
5		Item 15	.04	.07	-.10	*.17*																				
6		Item 17	-.02	-.03	-.13	.02	.10																			
7		Item 18	.14	*.19*	-.05	.09	.10	-.08																		
8		Item 2	.00	.04	.04	.07	.01	-.02	.04																	
9		Item 5	-.02	-.04	.02	-.01	.10	.01	.01	-.01																
10	Flexibility	Item 8	-.05	.04	-.03	-.12	.12	-.08	.04	.00	.05															
11		Item 13	.04	-.05	.01	.10	-.02	-.03	-.09	-.01	.00	-.01														
12		Item 16	.13	.11	.00	.03	.06	.03	*.18*	.13	-.11	.10	-.06													
13		Item 20	.00	.11	-.05	.09	-.02	.01	-.05	.12	.01	*.15*	-.07	-.03												
14		Item 21	.01	-*.15*	.03	.06	.06	*.15*	.11	.05	*.17*	.06	-.07	.10	*.16*											
15		Item 3	.08	.08	-.10	.12	.09	-.09	.10	.01	.08	.02	-.03	-.08	.01	.11										
16		Item 4	-.05	.06	.04	.09	.10	-.05	.12	-.04	*.17*	.14	.02	-.01	-.06	.03	.07									
17	Humility	Item 7	.04	.00	-.08	.08	-.02	.02	.03	-.07	.05	.03	.06	.00	-.03	.08	-.05	-.09								
18		Item 9	-.05	.04	.00	.06	-.04	.13	.07	-.05	-.05	-*.17*	.00	-.03	.01	.05	-.08	.06	*.17*							
19		Item 10	-.01	.03	.02	.07	.07	.08	-.03	.06	.05	-*.15*	***.23***	-.05	-.05	-.12	-.05	-.04	*.15*	.10						
20		Item 12	.00	-.03	-.01	.13	.03	.07	-.02	.09	.09	-.05	.00	.02	.01	***.21***	.07	-.12	.14	.02	-.04					
21		Item 19	.02	-.07	.02	.08	.05	.09	*.16*	-.03	-.08	.15	-.02	-.06	.13	.11	*.15*	.14	-.04	-.02	-.05	-.13				
22		Inclusiveness	***.43***	***.45***	*.16*	***.50***	***.51***	***.33***	***.48***	.06	.03	-.03	-.01	*.18*	.03	.11	.10	.11	.03	.08	.08	.07	.13			
23		Flexibility	.03	.02	.01	.07	.11	.03	.08	***.42***	***.43***	***.45***	***.22***	***.37***	***.49***	***.53***	.05	.09	.04	-.07	-.02	.13	.07	.13		
24		Humility	.00	.05	-.03	***.22***	.10	.09	*.17*	-.02	.11	-.01	.09	-.08	.02	*.16*	***.43***	***.44***	***.35***	***.48***	***.37***	***.25***	***.43***	***.22***	0.09	
25		Total Mindset	***.24***	***.26***	.08	***.40***	***.37***	***.23***	***.37***	***.25***	***.30***	***.22***	*.15*	***.25***	***.28***	***.41***	***.29***	***.32***	***.20***	***.23***	***.21***	***.22***	***.31***	***.69***	***0.64***	***0.65***

Note. ***Bold italics text:*** Correlation is significant at the 0.01 level (2-tailed). *Italics text*: Correlation is significant at the 0.05 level (2-tailed).

Table 11

Item-Scale and Validity Correlations (All Items)

		M	*SD*	*N*	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*
Inclusiveness	*1* Item 1	0.90	1.0	123									
	2 Item 6	0.88	0.9	123	0.03								
	3 Item 11	-0.41	0.9	123	-0.09	0.02							
	4 Item 14	0.41	1.1	123	0.07	0.12	-0.02						
	5 Item 15	0.63	1.1	123	0.05	0.09	-0.12	*0.20*					
	6 Item 17	0.41	1.2	123	-0.08	-0.07	-0.13	-0.06	0.13				
	7 Item 18	0.33	1.1	123	0.06	0.16	-0.08	0.15	0.06	-0.15			
	8 Inclusiveness	3.17	2.9	123	*0.37*	*0.44*	0.16	*0.55*	*0.53*	*0.28*	*0.44*	(.12)	
	9 Work-Related Openness	4.26	0.6	123	0.04	0.12	-0.07	0.07	0.15	0.08	0.15	*0.20*	(.89)
Flexibility	*1* Item 2	0.57	1.0	123									
	2 Item 5	0.71	1.2	123	0.00								
	3 Item 8	0.59	0.9	123	0.00	0.09							
	4 Item 13	0.59	0.9	123	-0.06	-0.11	-0.02						
	5 Item 16	0.69	1.0	123	0.13	-0.03	0.06	0.04					
	6 Item 20	0.18	1.1	123	0.09	0.02	0.17	-0.08	0.02				
	7 Item 21	0.59	1.1	123	0.00	0.14	0.01	-0.08	0.10	0.17			
	8 Flexibility	3.93	3.0	123	*0.39*	*0.44*	*0.43*	*0.20*	*0.44*	*0.50*	*0.48*	(.19)	
	9 Resistance to Change	3.16	0.7	123	-0.01	*-0.25*	-0.15	0.00	-0.03	*-0.19*	-0.11	*-0.27*	(.84)
Humility	*1* Item 3	0.25	1.1	123									
	2 Item 4	-0.28	1.2	123	0.11								
	3 Item 7	1.59	0.7	123	-0.05	-0.09							
	4 Item 9	0.51	1.2	123	-0.07	0.13	*0.24*						
	5 Item 10	0.41	1.1	123	-0.04	-0.06	0.15	0.10					
	6 Item 12	0.96	0.8	123	0.03	-0.12	0.10	0.02	-0.11				
	7 Item 19	0.03	1.1	123	0.17	*0.27*	-0.11	0.03	-0.03	-0.13			
	8 Humility	3.47	3.0	123	*0.44*	*0.51*	*0.32*	*0.53*	*0.35*	*0.18*	*0.47*	(.18)	
	9 Dispositional Humility	5.24	0.8	95	-0.01	0.07	*0.20*	*0.30*	*0.25*	-0.03	0.08	*0.31*	(.96)

Note. ***Bold italics text:*** Correlation is significant at the 0.01 level (2-tailed). *Italics text*: Correlation is significant at the 0.05 level (2-tailed).

Table 12

Item-Scale and Validity Correlations (Final Scale Items)

		M	*SD*	*N*	*1*	*2*	*3*	*4*	*5*	*6*
Inclusiveness	*1* Item 6	0.88	0.9	123						
	2 Item 14	0.41	1.1	123	0.12					
	3 Item 15	0.63	1.1	123	0.09	*0.20*				
	4 Item 18	0.33	1.1	123	0.16	0.15	0.06			
	5 Inclusiveness	2.26	2.5	123	***0.53***	***0.65***	***0.58***	***0.59***	(.37)	
	6 Work-Related Openness	4.26	0.6	123	0.12	0.07	0.15	0.15	*0.21*	(.89)
Flexibility	*1* Item 5	0.71	1.2	123						
	2 Item 8	0.59	0.9	123	0.09					
	3 Item 20	0.18	1.1	123	0.02	0.17				
	4 Item 21	0.59	1.1	123	0.14	0.01	0.17			
	5 Flexibility	2.07	2.5	123	***0.59***	***0.51***	***0.60***	***0.58***	(.31)	
	6 Resistance to Change	3.16	0.7	123	***-0.25***	-0.15	*-0.19*	-0.11	***-0.31***	(.84)
Humility	*1* Item 4	-0.28	1.2	123						
	2 Item 7	1.59	0.7	123	-0.09					
	3 Item 9	0.51	1.2	123	0.13	***0.24***				
	4 Item 10	0.41	1.1	123	-0.06	0.15	0.10			
	5 Humility	2.23	2.3	123	***0.52***	***0.45***	***0.70***	***0.52***	(.23)	
	6 Dispositional Humility	5.24	0.8	95	0.07	*0.20*	***0.30***	*0.25*	***0.36***	(.96)

Note: ***Bold italics text:*** Correlation is significant at the 0.01 level (2-tailed). *Italics text*: Correlation is significant at the 0.05 level (2-tailed).

Considering that each item score consists of a combination of an endorsement score (−1, 0, or 1) and a rejection score, it may be revealing to examine whether these scores, in isolation, have differential relationships with each other and other variables. To that end, Tables 13 and 14 present the validity correlations with the full combined item scores, as well as scores based only on endorsement or rejection. In Tables 15 and 16, the correlations between the endorsement-only scores and rejection-only scores are provided, at the item level and scale level.

Table 13

Group 1 Scale Validity Correlations (Final Scale Items)

		M	*SD*	*N*	*1*	*2*	*3*	*4*	*5*	*6*	*7*
Combined	*1* Inclusiveness	2.26	2.5	123	(.37)						
	2 Flexibility	2.07	2.5	123	0.10	(.31)					
	3 Humility	2.23	2.3	123	0.13	-0.01	(.23)				
	4 Total Mindset	6.56	4.5	123	***0.68***	***0.60***	***0.58***	(.36)			
	5 Dispositional Humility	5.24	0.8	95	0.00	*0.23*	***0.36***	***0.32***	(.96)		
	6 Resistance to Change	3.16	0.7	123	-0.06	***-0.31***	-0.17	***-0.29***	***-0.34***	(.84)	
	7 Work-Related Openness	4.26	0.6	123	*0.21*	0.14	*0.20*	***0.29***	***0.34***	***-0.51***	(.89)
Endorsement only	*1* Inclusiveness (End)	1.25	1.5	123	(.29)						
	2 Flexibility (End)	0.90	1.4	123	0.02	(.18)					
	3 Humility (End)	1.27	1.1	123	-0.02	-0.12	(.10)				
	4 Total Mindset (End)	3.42	2.3	123	***0.65***	***0.59***	***0.40***	(.11)			
	5 Dispositional Humility	5.24	0.8	95	-0.08	0.02	*0.26*	0.09	(.96)		
	6 Resistance to Change	3.16	0.7	123	-0.03	*-0.18*	-0.09	*-0.18*	***-0.34***	(.84)	
	7 Work-Related Openness	4.26	0.6	123	0.12	0.02	0.09	0.13	***0.34***	***-0.51***	(.89)
Rejection only	*1* Inclusiveness (Rej)	1.01	1.5	123	(.29)						
	2 Flexibility (Rej)	1.17	1.6	123	0.08	(.29)					
	3 Humility (Rej)	0.96	1.6	123	0.09	0.08	(.21)				
	4 Total Mindset (Rej)	3.14	2.9	123	***0.61***	***0.62***	***0.64***	(.34)			
	5 Dispositional Humility	5.24	0.8	95	0.07	***0.34***	***0.34***	***0.41***	(.96)		
	6 Resistance to Change	3.16	0.7	123	-0.07	***-0.33***	*-0.18*	***-0.31***	***-0.34***	(.84)	
	7 Work-Related Openness	4.26	0.6	123	*0.23*	*0.21*	*0.23*	***0.35***	***0.34***	***-0.51***	(.89)

Note: ***Bold italics text:*** Correlation is significant at the 0.01 level (2-tailed). *Italics text*: Correlation is significant at the 0.05 level (2-tailed).

Table 14

Group 2 Scale Validity Correlations (Final Scale Items)

			M	*SD*	*N*	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*
Combined	*1*	Inclusiveness	2.20	2.3	54	(.28)							
	2	Flexibility	2.35	2.6	54	0.02	(.31)						
	3	Humility	2.17	1.9	54	0.22	0.03	(.01)					
	4	Total Mindset	6.72	4.3	54	***0.65***	***0.63***	***0.59***	(.31)				
	5	Cognitive Ability	0.63	0.2	53	-0.01	-0.17	-0.05	-0.13	(.79)			
	6	STQ Reframing	3.54	0.5	54	-0.03	0.15	-0.12	0.02	-0.04	(.53)		
	7	STQ Systems Thinking	3.86	0.5	54	0.15	-0.07	-0.10	-0.01	-0.24	-0.03	(.62)	
	8	STQ Reflection	4.08	0.5	54	0.27	0.09	0.01	0.21	-0.14	***0.50***	***0.56***	(.60)
Endorsement only	*1*	Inclusiveness (End.)	1.24	1.4	54	(.20)							
	2	Flexibility (End.)	1.09	1.5	54	0.04	(.25)						
	3	Humility (End.)	1.31	1.0	54	0.14	0.01	(-.12)					
	4	Total Mindset (End.)	3.65	2.4	54	***0.66***	***0.66***	***0.50***	(.23)				
	5	Cognitive Ability	0.63	0.2	53	0.05	-0.26	-0.06	-0.16	(.79)			
	6	STQ Reframing	3.54	0.5	54	0.01	0.17	-0.04	0.10	-0.04	(.53)		
	7	STQ Systems Thinking	3.86	0.5	54	0.08	0.03	-0.13	0.00	-0.24	-0.03	(.62)	
	8	STQ Reflection	4.08	0.5	54	*0.29*	0.14	0.05	*0.28*	-0.14	***0.50***	***0.56***	(.60)
Rejection only	1	Inclusiveness (Rej.)	0.96	1.4	54	(.12)							
	2	Flexibility (Rej.)	1.26	1.5	54	-0.02	(.16)						
	3	Humility (Rej.)	-0.74	1.3	54	0.02	0.17	(-.05)					
	4	Total Mindset (Rej.)	1.48	2.6	54	***0.53***	***0.66***	***0.63***	(.17)				
	5	Cognitive Ability	0.63	0.2	53	-0.07	-0.03	-0.02	-0.07	(.79)			
	6	STQ Reframing	3.54	0.5	54	-0.06	0.08	*-0.31*	-0.15	-0.04	(.53)		
	7	STQ Systems Thinking	3.86	0.5	54	0.17	-0.14	-0.02	0.00	-0.24	-0.03	(.62)	
	8	STQ Reflection	4.08	0.5	54	0.16	0.01	-0.17	0.01	-0.14	***0.50***	***0.56***	(.60)

Note: ***Bold italics text:*** Correlation is significant at the 0.01 level (2-tailed). *Italics text*: Correlation is significant at the 0.05 level (2-tailed).

Table 15. *Item-Scale Rejection-Endorsement Correlations (Final Scale Items)*

			M	*SD*	*N*	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*
Inclusiveness	*1*	Item 6 (End.)	0.58	0.6	177										
	2	Item 14 (End.)	0.08	0.7	177	0.05									
	3	Item 15 (End.)	0.34	0.7	177	0.11	0.10								
	4	Item 18 (End.)	0.25	0.6	177	*0.17*	0.08	-0.01							
	5	Inclusiveness (End.)	1.25	1.4	177	***0.58***	***0.60***	***0.55***	***0.50***						
	6	Item 6 (Rej.)	0.29	0.5	177	***0.23***	-0.02	-0.04	0.11	0.12					
	7	Item 14 (Rej.)	0.27	0.7	177	0.11	***0.35***	*0.18*	0.10	***0.34***	0.03				
	8	Item 15 (Rej.)	0.28	0.7	177	0.07	0.05	***0.29***	0.03	***0.20***	0.02	0.13			
	9	Item 18 (Rej.)	0.16	0.7	177	*0.16*	0.09	0.03	***0.33***	***0.26***	0.04	-0.02	*0.19*		
	10	Inclusiveness (Rej.)	0.99	1.5	177	***0.25***	***0.22***	***0.22***	***0.26***	***0.42***	***0.42***	***0.52***	***0.64***	***0.60***	
Flexibility	*1*	Item 5 (End.)	0.39	0.6	177										
	2	Item 8 (End.)	0.25	0.6	177	0.01									
	3	Item 20 (End.)	0.12	0.8	177	-0.03	*0.17*								
	4	Item 21 (End.)	0.20	0.7	177	0.15	-0.04	0.10							
	5	Flexibility (End.)	0.96	1.5	177	***0.45***	***0.48***	***0.65***	***0.58***						
	6	Item 5 (Rej.)	0.34	0.8	177	***0.42***	-0.03	0.00	0.03	*0.16*					
	7	Item 8 (Rej.)	0.42	0.7	177	0.09	***0.24***	0.03	0.06	*0.18*	0.07				
	8	Item 20 (Rej.)	0.08	0.6	177	0.02	0.13	***0.26***	***0.24***	***0.32***	0.05	0.07			
	9	Item 21 (Rej.)	0.36	0.6	177	0.13	0.10	0.03	***0.31***	***0.26***	*0.19*	0.03	0.04		
	10	Flexibility (Rej.)	1.20	1.5	177	***0.33***	*0.18*	0.13	***0.27***	***0.40***	***0.67***	***0.51***	***0.48***	***0.54***	
Humility	*1*	Item 4 (End.)	0.01	0.5	177										
	2	Item 7 (End.)	0.84	0.4	177	-0.02									
	3	Item 9 (End.)	0.25	0.5	177	-0.06	0.06								
	4	Item 10 (End.)	0.18	0.7	177	-0.02	0.08	0.02							
	5	Humility (End.)	1.28	1.1	177	***0.43***	***0.44***	***0.48***	***0.65***						
	6	Item 4 (Rej.)	-0.27	0.9	177	***0.33***	0.01	0.09	-0.10	0.15					
	7	Item 7 (Rej.)	0.27	0.9	177	*-0.15*	0.14	0.09	0.06	0.06	-0.02				
	8	Item 9 (Rej.)	0.24	0.9	177	0.07	*0.16*	***0.41***	0.02	***0.30***	0.03	0.06			
	9	Item 10 (Rej.)	0.20	0.6	177	0.01	0.11	0.03	***0.27***	***0.23***	0.02	0.06	*0.17*		
	10	Humility (Rej.)	0.44	1.7	177	0.13	***0.20***	***0.31***	0.09	***0.33***	***0.51***	***0.54***	***0.61***	***0.48***	

Note. ***Bold italics text***: Correlation is significant at the 0.01 level (2-tailed). *Italics text*: Correlation is significant at the 0.05 level (2-tailed).

Table 16

Scale Level Correlations Between Rejection and Endorsement Scores (Final Scale Items)

		M	*SD*	N	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*
1	Inclusiveness (End.)	1.25	1.4	177								
2	Flexibility (End.)	0.96	1.5	177	0.03							
3	Humility (End.)	1.28	1.1	177	0.02	-0.08						
4	Total Mindset (End.)	5.50	3.1	177	***0.53***	***0.51***	***0.38***					
5	Inclusiveness (Rej.)	0.99	1.5	177	***0.42***	-0.07	*0.17*	***0.29***				
6	Flexibility (Rej.)	1.20	1.5	177	***0.19***	***0.40***	0.05	***0.32***	0.05			
7	Humility (Rej.)	0.44	1.7	177	0.09	-0.09	***0.33***	0.03	0.07	0.08		
8	Total Mindset (Rej.)	4.97	3.7	177	***0.33***	*0.16*	***0.31***	***0.40***	***0.53***	***0.51***	***0.49***	

Note: ***Bold italics text:*** Correlation is significant at the 0.01 level (2-tailed). *Italics text*: Correlation is significant at the 0.05 level (2-tailed).

Hypothesis 2 covered the expected relationships between the STMT subscales and the alternative measures of similar constructs to establish construct validity. Specifically, Hypothesis 2a predicted a negative relationship between strategic thinking mindset flexibility and resistance to change. The observed correlation between the four-item flexibility score and resistance to change was significant and negative, as predicted ($r = -.31$, $p < .01$). Hypothesis 2b predicted a positive relationship between strategic thinking mindset inclusiveness and work-related openness. The observed correlation between the four-item inclusiveness score and work-related openness was significant and positive, as predicted ($r = .21$, $p < .05$). Hypothesis 2c predicted a positive relationship between strategic thinking mindset humility with dispositional humility. The observed correlation between the four-item humility score and dispositional humility was significant and positive, as predicted ($r = .36$, $p < .01$). See Table 13 for the validity correlations in Group 1.

Using endorsement scores only, the flexibility-resistance to change correlation is significant (−.18, < .05), but smaller than the rejection-based correlation (−.33, $p < .01$). This effect is repeated for the other two scales as well. The endorsement-based correlation between inclusiveness and work-related openness (.12, *ns*) is smaller than the rejection-based (.23, $p < .05$), and the endorsement-based correlation between humility and dispositional humility (.26, $p < .05$) is smaller than the rejection-based correlation (.34, $p < .01$). Within the STMT, the endorsement and rejection scores for the inclusiveness and flexibility scales each correlate at around .4, while the two humility scores correlate a bit lower, at .33. In sum, the construct validity of the STMT, as indicated by its relationship with similar constructs, is based more on the response options that are rejected than the ones that are endorsed.

Focusing again on the combined scores (endorsement and rejection together), the relationships between each characteristic and the other construct validity measures should be explored. Inclusiveness did not correlate significantly with either resistance to change ($r = -.06$) or dispositional humility ($r < .01$). Flexibility did not correlate significantly with work-related openness (r = .14), but did correlate with dispositional humility ($r = .23$, $p < .05$). Humility did

not correlate with resistance to change (−.17), but did correlate significantly with work-related openness (.2, $p < .05$). Taken as a whole, this evidence cumulatively supports Hypotheses 2a–c, although the reliability estimates based on coefficient alpha suggest a need for further examination of the test's reliability.

In Group 2, additional measures were included to further explore the relationships between the STMT and other relevant variables, specifically the Pisapia et al. (2005) STQ and the ICAR cognitive ability test. All correlations between the STMT subscales and the STQ subscales were non-significant, further supporting the discriminant validity of the SJT, in that the constructs the STMT was intended to measure are different than the reflection, reframing, and systems thinking behaviors measured by the STQ. The observed correlation between the overall STMT score and cognitive ability was also non-significant, indicating that the STMT, as intended, was measuring constructs distinct from cognitive ability. See Table 14 for the validity correlations in Group 2.

Additional analysis revealed no demographic effects. There was no effect of gender on scores for inclusiveness ($F(1,175) = 0.22$, *ns*), flexibility ($F(1,175) = 0.33$, *ns*), humility ($F(1,175) = 1.07$, *ns*), or total mindset ($F(1,175) = 0.01$, *ns*) as measured by the STMT. Likewise, there was no effect of gender on work-related openness ($F(1,121) = 0.45$, *ns*), resistance to change ($F(1,121) = 0.06$, *ns*), dispositional humility ($F(1,93) = 1.06$, *ns*), reframing ($F(1,52) = 0.92$, *ns*), systems thinking ($F(1,52) = 0.17$, *ns*), reflection ($F(1,52) = 0.34$, *ns*), or cognitive ability scores ($F(1,52) = 2.15$, *ns*). It should be noted that, although the sample's balance of males and females is consistent with the Army-wide population (cf., Maxfield, 2015), there is a large imbalance in the sample for this demographic variable.

There was no effect of rank on scores for inclusiveness ($F(2,174) = 0.06$, *ns*), flexibility ($F(2,174) = 1.33$, *ns*), humility ($F(2,174) = 1.37$, *ns*), or total mindset ($F(2,174) = 1.25$, *ns*). There was a significant, but trivial, effect of rank on work-related openness ($F(2,120) = 5.8$, $p < .01$), 2LTs ($M = 3.66$) scored lower than CPTs ($M = 4.33$), but no effect of rank on resistance to change ($F(2,120)$, 1.49, *ns*) dispositional humility ($F(2,92) = 0.58$, *ns*), reframing ($F(2,51) = 1.07$, *ns*), systems thinking ($F(2,51) = 1.47$, *ns*), reflection ($F(2,51) = 2.6$, *ns*), or cognitive ability scores ($F(2,51) = 0.53$, *ns*).

There was no effect of functional branch on scores for inclusiveness ($F(16,160) = 1.15$, *ns*), flexibility ($F(16,160) = 0.33$, *ns*), humility ($F(16,160) = 1.1$, *ns*), or total mindset ($F(16,160) = 0.49$, *ns*). Likewise, there was no effect of functional branch on work-related openness ($F(14,108) = 1.64$, *ns*), resistance to change ($F(14,108) = 0.47$, *ns*), dispositional humility ($F(14,80) = 1.15$, *ns*), reflection ($F(10,43) = 1.71$, *ns*), or cognitive ability scores ($F(10,42) = 1.4$, *ns*). There were two significant, but trivial, effects of functional branch: reframing ($F(10,43) = 2.89$, $p < .01$), mean branch scores ranged from 3.2 to 4.0, and systems thinking ($F(10,43) = 2.8$, $p < .01$), mean branch scores ranged from 3.3 to 4.2.

Discussion

Gaining new understanding about the improvement of organizational strategy through the development of effective strategic thinkers is a goal shared by many academic scholars, human resource professionals, and organizational leaders. It has only been fairly recently that the so-

called black box of the boardroom, the often empirically inaccessible environment in which high-level strategizing happens in large organizations, has started to open to scholars and researchers (Arnardottir, Fischer, & Martin, 2015). Much remains to be known about the individual and interpersonal dynamics of strategic thinking and strategizing. The current research sought to contribute to this understanding by conceptualizing a new construct: the strategic thinking mindset. In conjunction with cognitive ability and the domain-specific knowledge best gained with time, experience, and education, such a mindset reflects the unique perspective that makes a person well-suited to the complex and ambiguous thinking challenges required for successful strategy-making.

The overarching goal of creating a test that measures the strategic thinking mindset was to help identify strengths related to strategic thinking in organizational members that have yet to be set on a path to strategic-level leadership. This research effort aimed to develop a theoretical understanding of the strategic thinking mindset as a construct, but also to measure that construct with a situational judgment test format conducive to identification and development in a way that is specific to an organization's work domain.

In brief, existing scholarship on the characteristics of strategic thinking was summarized and synthesized by an analysis of the themes that underlie the breadth of the construct across several disciplines. Once the characteristics of the strategic thinking mindset were identified, four stages of progressive item development and refinement produced a 12-item situational judgment test, the STMT, for use by the U.S. Army with three sub-scales that show evidence of construct validity, in line with the theoretical model.

The conduct of the research was structured in four stages. The first three stages of data collection and analysis reflect the development of the STMT items. The standard form of a situational judgment test item consists of an item stem (a scenario of some kind), item options (ways in which one could respond in the scenario), and a scoring key (how each response option relates to the item score, if chosen). The fourth stage consisted of the evaluation of the STMT's psychometric qualities.

In Stage 1, participants from a wide range of functional areas in the Army provided key incidents from their experiences in which they or a close associate were confronted with a difficult problem that had no clear right answer. These key incidents were carefully reviewed, culled, and edited into a set of scenarios that elicit a demonstration (or rejection) of humility, flexibility, or inclusiveness by a functionally generic audience of Army participants at a certain level of rank.

In Stage 2, a new group of participants read the scenarios and provided feedback on their realism and general applicability across the Army, as written. Participants also indicated the most relevant characteristic for the scenario and provided a range of response options, with emphasis on varying the focal characteristic within a range of reasonably realistic and effective approaches to responding to the problem depicted in the scenario. Scenario feedback was incorporated, as needed, to improve the scenarios. The votes for each scenario as relevant to flexibility, humility, or inclusiveness were tallied and used to assign each scenario a characteristic. The many response options provided for each scenario were reviewed and

compiled into a set of candidate response options that represented different levels of the focal characteristic.

In Stage 3, a different group of Army participants considered the options for each scenario and rated them individually on the level of flexibility, humility, or inclusiveness shown and the likely effectiveness of the option. A small group of research psychologists also rated the scenarios and options in the same way. These data were carefully compiled and analyzed in an attempt to select the most appropriate options for each scenario and identify the scenarios with the best group of response options to represent a range of flexible, humble, or inclusive responses, with generally equivalent effectiveness.

Finally, in Stage 4, the remaining scenarios and response options were presented to a final group of Army participants as pilot test items, alongside measures of other constructs for accumulating validity evidence. Validity evidence was gathered using measures of resistance to change, work-related openness, dispositional humility, reframing, reflection, systems thinking, and cognitive ability.

The results demonstrated that the STMT's low inter-item correlations and construct heterogeneity led to low internal consistency reliability estimates. This is not particularly surprising given the issues of construct heterogeneity inherent to the SJT format (see Whetzel & McDaniel, 2009). Although some evidence emerged supporting the theoretical model underlying the STMT's structure, the lack of an adequate indicator of the test's reliability and the questionable factor structure of the test may limit the test's usability as a predictor or outcome criterion.

If the STMT were to be employed in research or selection, a better estimate of the reliability of the test, such as a test-retest form, would be needed before the test could be trusted for this purpose. Any criterion-related correlations would need to be attenuated for unreliability in the test. The single test administration nature of the current research posed a difficult challenge to determining the test's true reliability. The construct heterogeneity of SJTs contributes to a lack of internal consistency of SJTs (Whetzel & McDaniel, 2009). Many other SJT developments have encountered similar issues. Schmitt and Chan (2006) describe the common occurrence of relatively low alphas, low inter-item correlations, and factor analytic results that account for a small portion of total variance in SJT development research. Porterfield (2001) developed an SJT for security officers, finding it to be an effective predictor of job performance, but suffering from a disappointing alpha (.31). Chan and Schmitt (1997) developed an SJT for generic skilled blue-collar work with a final alpha of .55.

There are, however, examples of SJTs that succeed in (or come close to) meeting the standard thresholds of alpha reliability (e.g., Ascalon, 2004, .69; Born, Van der Maessen, & Van der Zee 2001, .91). It has been suggested that an increased focus on measuring specific constructs, rather than the entire scope of a job, may lead to better alphas (Ascalon, 2004). Born et al. (2001) developed a construct-focused SJT for social intelligence and achieved a very high alpha; however, Young's (2004) SJT focused on emotional intelligence suffered from an extremely low alpha (.17). Likewise, an SJT developed by Smith (2011) that focused on entrepreneurial orientation produced an alpha of .32. Furthermore, coefficient alpha, as the

standard metric for internal consistency reliability, has been criticized as a (potentially gross) underestimate of true reliability (Sijtsma, 2009).

Schmitt (1996) notes that a measure may have strengths other than high internal consistency reliability around a clear homogeneous construct, such as content coverage of a domain, that may support the test's use. The STMT's reliability is of greater concern when the purpose of the test is to analyze the score. Yet, one of the great strengths of an SJT is its content fidelity (Weekley & Ployhart, 2006). The use of the STMT as a starting point for self-awareness training and group discussion about intellectual flexibility, humility, and inclusiveness is not substantially hindered by the unknown reliability and factor structure of the test. The results regarding reliability are disappointing from a pure scale construction perspective. However, this outcome is consistent with the issues that have hindered many previous SJT developments.

There is some evidence that the measure taps into the constructs of flexibility, humility, and inclusiveness as intended. In support of Hypotheses 2a–c, each three-item subscale in the final test correlated as expected with its corresponding self-report measure, but did not correlate significantly with the others, aside from the STMT humility scale correlating with work-related openness. While there may be notable construct overlap and statistical noise in each subscale's strength as a predictor and construct measurement, there is significant scale-level variance uniquely tied to the intended construct. Exploration of the endorsement-only and rejection-only scores indicated that a participant's rejection choice contributed more significantly to the variance that was shared with the construct validity scale, indicating that there may be greater validity in what a leader knows *not* to do, in regard to their intellectual inclusiveness, intellectual flexibility, and intellectual humility.

That said, there is a reason why reliability, inter-item and item-total correlations, and factor structure are a concern in measurement, particularly when a test is designed to measure specific constructs, rather than broadly defined "judgment" in a job. In the current research, the STMT was designed to measure the construct of strategic thinking mindset as indicated by the constructs of flexibility, inclusiveness, and humility. There remain important concerns about the overall evidence supporting the construct validity of the STMT in this regard. The low inter-item correlations suggest that each item reflects a heterogeneous set of constructs, even within the subscales.

The results related to cognitive ability are noteworthy, in that they reinforce the constructs being measured by the STMT as distinct from a general intelligence factor, as intended. They also reinforce the need to appreciate and understand the aspects of strategic thinking that require something other than a powerful intellect. In fact, highly intelligent people might have a tendency, either innate or learned, to devalue certain ways of approaching problems. Many may tend to disregard the viewpoints of others (inclusiveness), fail to acknowledge personal limits or biases (humility), or disregard clues that a change in approach or assumptions may be needed (flexibility). More research on this possibility is warranted.

The lack of relationships between the STMT and STQ (Pisapia et al., 2005), aligns with the different approach to conceptualizing the strategic thinking mindset in each measure. The approach taken by the STQ focuses more on three "cognitive processes" (p. 44): systems thinking, reflection, and reframing. According to Pisapia et al. (2005), these processes facilitate

the creation of the strategic thinking mindset. These are certainly important elements of strategic thinking, as indicated by the inclusion of the Pisapia et al. model in the common themes analysis summarized in Table 1. The STMT focuses on the mindset characteristics that form a necessary foundation for the kinds of thinking processes measured in the STQ. Additionally, it should be noted that the methodological differences of the STMT and STQ as measures affect what exactly is being measured. While the STQ relies on the respondent's self-assessment of general tendencies, the STMT puts the respondent in a position to apply those tendencies in a series of situational judgments. The significant relationships found between inclusiveness and the total mindset score with reflection may suggest an avenue for further research. For example, it may be that inclusiveness plays a role in supporting the processes represented by the theme of self-awareness and self-control.

Strengths of the Research

Although the overall profile and strength of the results for the scale developed in this research are somewhat disappointing, there are still important strengths in the areas of conceptual development, method selection and item development, and research design. The conceptualization of the strategic thinking mindset may still need further refinement, but the process by which the mindset was conceptualized was sound, rooted in a diverse review and synthesis of existing scholarship from the disciplines of psychology, management, military art, and history. There is far too much empirical and conceptual literature about strategy to provide a fully comprehensive review of the field. The scope for this research was to examine how different disciplines ascribe specific characteristics to individual strategic thinking. In some cases, this included explicitly designed scientific models of strategic thinking, while in other cases, the substance of strategic thinking is described in a more philosophical way, or embedded in an approach to measurement. This is a strength of this research, in that it avoids what can become a self-reinforcing cycle of perpetually winnowed concepts relying too heavily on one contextual lens. The notion of a strategy is unquestionably broad and fundamental to individuals and organizations of all types, be they small or big, private or public, military or civilian, athletic or artistic. There is a cost to the broader understanding of how to think about a complex problem environment and create effective strategies within it if researchers and scholars fail to maintain a sufficiently wide lens.

This research intentionally sought to integrate the common themes of all who think and write about what strategic thinking is and looks like, and then to take a step further in using that synthesized understanding to derive a new concept. For all that we might already understand about what strategic thinking is and who does it well, we know much less about how to identify those who have a talent for it and give them the right developmental experiences, such as planning and decision-making opportunities, broadening experiences, and education. The mindset for strategic thinking, how to approach complex problem environments in the right way, is an important tool, alongside intelligence and a base of relevant knowledge, for individuals to contribute to effective strategic thinking and strategy-making.

The choice of situational judgment testing as a method for assessing the strategic thinking mindset was thoughtfully rooted in the strengths of this format. First and foremost, the SJT format is a relatively inexpensive method that retains some content fidelity as a small-scale simulation, rather than relying on a test-taker's self-awareness and honesty. Face validity to test-

takers, reduced adverse impact, and potential material for group-discussion and self-development also weighed in favor of this method. The item development process was carefully planned and performed in line with the best practice guidance of SJT experts and practitioners, notably McDaniel and Whetzel (2007). Each element of the STMT (item stem content, item response content, scoring key) was developed in its own stage, allowing for each stage of development to contribute to improving the products of the previous stages.

At each stage, the STMT was developed with a unique and relevant sample of participants with expertise in the job environment at hand. Employing a more generic student sample may have facilitated easier data collection and larger samples, which may have facilitated greater flexibility for maximizing the psychometrics. However, the utility of the STMT, in the end, would have suffered significantly from the lack of perspective on the real-life patterns of thought and decision-making in the population for whom the test was intended.

Limitations of the Research

The primary limitation of this research relates to the lack of a solid understanding of what the STMT is truly measuring at a construct level, as indicated by the low inter-item correlations. The reliability problems are inherently related, in that they rely on the internal consistency model of estimating reliability, which treats each item as a mini-test; therefore, the degree to which the items intercorrelate with each other determines the estimate of how reliable the test is. As noted above, and according the recommendation of Whetzel and McDaniel (2009), this approach to estimating the reliability of an SJT is suboptimal compared to other approaches, such as test-retest reliability.

As noted above, SJTs commonly present muddled results in regard to identifying the constructs measured; the STMT was no different. The nature of the Army sample should be considered a strength of this research overall. However, there are limitations aside from the restrictions on sample size and restraints on the opportunity to estimate reliability through test-retest correlations. Although the samples obtained certainly have expertise derived from first-hand knowledge of decision-making in the job, there is no real expertise in the sample regarding the constructs of flexibility, humility, and inclusiveness. Although the author's involvement at all points, and the research psychologist sample in Stage 3, serve to balance this concern, the participants may have been limited in their ability to sufficiently conceptualize the relevance of a scenario to the mindset characteristics or to generate a wide enough range of realistic flexible, humble, and inclusive response options.

There are also limitations to the strength of the strategic thinking mindset at a conceptual level. As noted throughout the introduction, there is likely a curvilinear relationship between each of the strategic thinking mindset characteristics and success in strategic thinking (on the curvilinear relationship between traits and performance, see Le et al., 2011). In other words, there is reason to believe that, beyond a certain threshold, these characteristics may become a hindrance to strategic thinking. This research does not allow for an investigation of that phenomenon; however, it bears mentioning here as the application of these characteristics in problem-solving requires a judgment about where that threshold may be. Some respondents may value inclusiveness highly, but feel that in the scenario presented, to embrace an inclusive approach to the problem would be more problematic than helpful.

A related point regarding the interpretation of one's mindset from SJT responses is that one's natural tendencies may be suppressed by the wishes of an authority figure or the norms of a culture. This is of particular relevance in the military, which has a clearly defined chain of command, powerful cultural and climate norms, and constant peer competition for promotion. A respondent may be naturally inclined to endorse an inclusive response to a given scenario, but also feel that such a response would be perceived badly by peers or authority figures. For example, an inclusive response may be seen as an abdication of one's own decision-making responsibility, perhaps with the intent of sharing the risks associated with failure.

In this case, it would be more accurate to say that the STMT is measuring the degree to which the unit climate a respondent comes from is receptive to flexibility, humility, and inclusiveness. Future research could effectively explore this phenomenon by comparing the results of Soldiers within and between units. This also points back to the measure's strength as a source of group discussion and self-development. If unit leaders wish to get a sense of how their command climate does or does not support flexibility, inclusiveness, and humility, the test could be used as a group diagnostic and source of material for discussion.

Future Research

As noted above, future research for the STMT, in particular, should focus on the question of whether it tends to measure an innate trait or a more transient perception of the characteristics, perhaps brought on by climate and/or leadership, using a between- and within-unit design. Although there were no significant effects of branch on scores, a more directed effort to obtain a sufficient sample from each of the major branches may reveal something about the areas in the Army that tend to produce (or attract) individuals with a strategic thinking mindset, and might, therefore, be under- or over-represented in promotion to strategic-level leadership positions. If the mindset is more transient and influenced by unit, this would suggest that the mindset could be developed through education or facilitated by improvements to unit climates. This is a critical area of future research on the mindset as it stands to define whether, and through what time and resources, any Soldier can be trained to think this way, or if organizations are better off facilitating the promotion of Soldiers who already think about problems in this way.

Related to the issue of transience is the question of how one's mindset relates to one's actions. The SJT format uses decisions about actions to indicate tendencies about a test-taker's mindset. Therefore, it is conceptually possible that a person who does not have a mindset favoring those characteristics would still make decisions and act in a way that suggests they do. This might happen, for example, if the Army's evaluation and promotion processes begin to explicitly evaluate and reward these characteristics. Future research could explore this possibility, perhaps through an experimental design that involves priming a group of participants that the Army highly values those characteristics. It may not matter if inclusiveness, humility, or flexibility are demonstrated in their actions, but not in their thinking. For example, subordinates only witness the actions and may, therefore, learn to appreciate the role of these characteristics in problem solving. However, when it comes to the individual officer's ability to make good decisions, if they only go through the motions of inclusiveness, for example, without truly incorporating the perspectives provided, their thinking will still suffer from a narrow lens.

The STMT should be tested in an environment in which test-retest reliability can be estimated, as well as other methods for evaluating the validity of the test, such as criterion-related validity. A research design might utilize course performance at various Army schools, such as Captain's Career Courses or the Command and General Staff College. Depending on the nature of the course, this may be a useful performance criterion for the test. Conversely, the STMT could be used to evaluate training interventions related to flexibility, humility, or inclusiveness.

The relationship of the STMT with cognitive ability could also be explored more fully. Although the results supported the independence of the strategic thinking mindset and cognitive ability, a qualitative and quantitative approach could be designed to explore how people high and low in cognitive ability tend to think about and value the importance of flexibility, humility, and inclusiveness. The STMT's relationship with personality and other individual difference variables should also be explored. Although one element of the five-factor model was incorporated in this research (the WROS focuses on the openness to experience factor), there may be significant overlap with one or more of the other four factors. Need for cognitive structure may be revealing of why certain individuals struggle to embrace a mindset of flexibility, for example. All such analyses would likely help to clarify the constructs being measured in the STMT.

Finally, further research could be done to clarify the relationship of the strategic thinking mindset to problem-solving and strategic thinking ability. For example, how does domain-relevant knowledge affect the strategic thinking mindset? It may be that more domain-relevant knowledge decreases the ability of the strategic thinking mindset to influence decision-making, as the individual might use this domain knowledge to focus more on the particular facts of the situation. If this were the case, it may be preferable to measure the mindset when officers are still recruits and have yet to gather any knowledge or experience at the Company grade. However, a recruit may be motivated to attain high scores on flexibility, humility, and inclusiveness when presented with scenarios that they know are beyond their training, education, and experience. When the scenarios match the level of the test-takers, however, higher mindset scores may be more indicative of how they would actually think and behave. Finally, as shown in Figure 2, two of the common themes were not tied to a mindset characteristic: directionality and awareness of time. Future research should seek to develop measures of those processes and assess whether they relate to the strategic thinking mindset.

Other approaches to measuring the strategic thinking mindset should also be explored. Some SJTs are developed using multi-media formats, depicting the scenario through audio-visual means, for example. This presents a substantial resource challenge; however, a multi-media version of the assessment might allow for an investigation of how interpersonal biases (racial, ethnic, gender, etc.) or stereotypes affect scores on the STMT. For example, someone biased against female officers in a combat environment may be much less likely to be inclusive than if a male officer is presented in the same role.

A self-report measure similar to the STQ (Pisapia et al., 2005) might be developed, based on the specific concepts of flexibility, humility, and inclusiveness in problem-solving as modeled in this research. This would allow for a better comparison of the way in which people assess their own tendencies and how they apply them in scenario-based judgments. The issue of social

desirability in responding is certainly relevant here. No explicit measure of socially desirable responding was included in this research; however, future studies should examine this issue, both for the SJT and especially if a self-report version of the STMT is developed. The STMT should be considerably more resistant to socially desirable responding, as the constructs being measured in each item are more obscured and embedded than, for example, in the STQ. That said, future research could explore how honest or accurate participants' responses to the SJT are in relation to how they actually behave in problem-solving. An observational study could be designed in which participants take the SJT and participate in one or more group problem-solving exercises.

A qualitative approach might also prove valuable. The growing field of computational linguistics and natural language processing might offer new options for assessing a person's cognitions in their written or oral communication. Critical self-reflection essays, lessons-learned reports, or after-action review transcripts may contain sufficient content to assess the mindset. Similarly, such an approach could be used to further refine the strategic thinking mindset concept itself by obtaining written or oral reflections on decision making by individuals in strategic positions and mining the ways and frequency with which they refer to moments of flexibility, inclusiveness, and humility in their thought process.

Future research in the area of SJT development in general should focus on keying and scoring methods (e.g., Bergman, Drasgow, Donovan, Henning, & Juraska, 2006). Further refining the circumstances in which each of the many options is most desirable and defining the ways in which choices in keying and scoring impact the data will help future researchers better anticipate and account for the data-related problems so commonly faced by SJTs. Specific research in this line might focus on alternate methods of keying and scoring the STMT in particular: altering the nature of the keying system employed to explore the effect of those changes on relationships with variables in the current or future studies. One example might be to weight each item in the total scale score according to its discrepancy between the positive and negative keyed options as scored in Stage 3. This would give greater weight to scores of items that feature a clearer delineation of the characteristic among the options.

Finally, future SJT research should seek to expand the SJT as a computer-adaptive test. Here, the test taker might experience a choose-your-own-adventure test wherein each response option to an initial scenario would have its own item characteristics based on an item response theory analysis. As with any computer-adaptive test, the test taker's response would drive the next test item presented. Instead of an entirely separate item presented next, the next item could be a new scenario depicting the same character confronted with the consequences of the initial choice, with a set of response options that have item characteristic curves that will help refine the estimate of the test-taker's true ability level.

There are obvious practical challenges to developing the content for such a test. However, this approach would be a natural fit for the assessment of strategic thinking, as it would allow a test-taker to experience the evolution of a problem and how his or her choices can shape the environment. An even loftier goal might be to develop a system that allows for free responding to each scenario, rather than a multiple-choice format. If an adequate algorithm could be derived through computational linguistics to evaluate the response in relation to a bank of existing coded responses, the item could be scored the same as any other SJT, but without providing the range of possible answers for the respondent.

Implications

Despite the somewhat unclear psychometric properties of the scale developed, there are useful implications that can be derived from this research. The primary implication for the theory of strategic thinking is the ability to put to use the common themes. As shown in Table 1, the various models/descriptions of strategic thinking published in articles and books from various disciplines cover much of the same content domain. Furthermore, many of the same terms are often used. However, each of the models leaves out at least a few of the important concepts described in the other models and, in some cases, slightly different terms are used to describe the same general idea. Although each scholar has the right to his or her own unique understanding and preferred terminology, this can create problems for the practical application of this scholarship to the development of strategy and strategic thinking. The nine themes provide a description of the shared space that theorists in fundamentally different environments have described when identifying what strategic thinking is and what it requires. Although the environments of strategizing for private-sector competitive position in industrial markets and grand military strategy for the use of the elements of national power differ widely in what their strategies consist of and look like, the art and science of strategic thinking are largely the same. Practitioners and theorists interested in developing, identifying, coaching, or evaluating strategic thinkers can use these themes to ensure they are operating with all the valuable insights brought by scholars in various disciplines.

In addition, the strategic thinking mindset conceptualized and pursued in this research has similar implications. Further research must be done to refine the understanding of what the mindset consists of and how it is best measured and applied in problem-solving. The general description of the mindset, distinct from intelligence and knowledge, provides an avenue for organizational solutions related to producing better strategic thinkers in organizations. Ensuring that organizational climates are maintained that support the growth of the mindset, and do not suppress or eliminate it, is an important implication of this concept. Likewise, organizations would benefit from providing opportunities to apply the mindset in practice, particularly among those with exceptional intelligence and high levels of the mindset characteristics who have yet to be put in the position to engage with strategic level problem-solving.

Conclusion

The current research was designed to accomplish three main objectives: (a) to build a theoretically sound concept of a strategic thinking mindset that is rooted in the understanding of strategic thinking in multiple disciplines, (b) to develop a test of the mindset for U.S. Army officers using situational judgment testing, and (c) to evaluate the psychometric and construct validity evidence of the test.

Although the objective to create and evaluate such a test was accomplished, the implicit objective that the test successfully demonstrate psychometric utility and validity is much more ambitious and presented an elevated risk of failure. Strategic thinking is a highly subjective phenomenon that is challenging to define and to operationalize in psychological measurement. Furthermore, the SJT format was chosen to assess the strategic thinking mindset for its potential as a middle ground solution between a high-fidelity simulation exercise technique and a behavioral self-report technique reliant on the honest reflection and self-awareness of a test-

taker. However, the SJT format brings with it significant issues for constructing a reliable and valid scale of identifiable constructs.

While the theoretical work in integrating theories of strategic thinking and conceptualizing the mindset construct represent useful contributions to theory and practice, the evidence supporting the psychometric quality of the STMT and its construct-related validity may need further exploration. There is some evidence suggesting that the STMT measures what it was designed to measure; however, it is unclear how reliable the results are on an individual basis. Additional research on the STMT, other methods of measuring the strategic thinking mindset, and further clarification of the concept of a strategic thinking mindset are needed.

References

Abraham, S. (2005). Stretching strategic thinking. *Strategy & Leadership, 33*(5), 5–12. doi: 10.1108/10878505106168 34

Alberts, D. S., & Hayes, R. E. (2003). *Power to the edge: Command and control in the information age*. Washington, DC: U.S. Department of Defense Command and Control Research Program.

Anderson, P. (1999). Complexity theory and organization science. *Organization Science, 10*(3), 216–232. doi: 1047-7039/99/1003/0216

Argyris, C. (1991). Teaching smart people how to learn. *Harvard Business Review, 69*(3), 99–109.

Arnardottir, A. A., Fischer, M. D., & Martin, T. A. (2015). Inside the black box: Board observation in practice. *Academy of Management Proceedings* (Meeting Abstract Supplement). doi: 10.5465/AMBPP.2015.11810symposium

Arthur, W., & Villado, A. J. (2008). The importance of distinguishing between constructs and methods when comparing predictors in personnel selection research and practice. *Journal of Applied Psychology, 93*, 435–442. doi: 10.1037/0021-9010.93.2.435

Ascalon, M. E. (2004). *Improving expatriate selection: Development of a situational judgment test to measure cross-cultural social intelligence*. (Doctoral Dissertation). Retrieved from ProQuest Dissertations & Theses Global. (3148748).

Barling, J., Christie, A., & Hoption, C. (2011). Leadership. In S. Zedeck (Ed.), *APA Handbook of Industrial and Organizational Psychology, Vol. 1: Building and Developing the Organization* (pp. 183-240). Washington, DC: American Psychological Association.

Bergman, M. E., Drasgow, F., Donovan, M. A., Henning, J. B., & Juraska, S. E. (2006). Scoring situational judgment tests: Once you get the data, your troubles begin. *International Journal of Selection and Assessment, 14*(3), 223-235. doi: 10.1111/j.1468-2389.2006.00345.x

Bethel, S. A., Prupas, A., Ruby, T. Z., & Smith, M. V. (2010). Developing Air Force strategists: Change culture, reverse careerism. *Joint Forces Quarterly, 58*(3), 82–88.

Bonn, I. (2001). Developing strategic thinking as a core competency. *Management Decision, 39*(1), 63–70. doi: 10.1108/EUM0000000005408

Bonn, I. (2005). Improving strategic thinking: A multilevel approach. *Leadership and Organizational Development Journal, 26*(5), 336–354. doi: 10.1108/01437730510607844

Born, M. P., Van der Maessen, P. E. A. M., & Van der Zee, K. I. (2001). A multimedia situational judgment test for social intelligence. *Journal of eCommerce and Psychology, 2*(2), 54–72.

Budner, S. (1962). Intolerance of ambiguity as a personality variable. *Journal of Personality, 30*(1), 29–50. doi: j.1467-6494.1962.tb02303.x

Casey, A. J., & Goldman, E. F. (2010). Enhancing the ability to think strategically: A learning model. *Management Learning, 41*(2), 167–185. doi: 10.1177/1350507609355497

Chan, D., & Schmitt, N. (1997). Video-based versus paper-and-pencil method of assessment in situational judgment tests: Subgroup differences in test performance and face validity perceptions. *Journal of Applied Psychology, 82*(1), 143–159.

Cherniss, C., & Goleman, D. (2001). *The emotionally intelligent workplace*. San Francisco, CA: Jossey-Bass.

Condon, D. M., & Revelle, W. (2014). The international cognitive ability resource: Development and initial validation of a public-domain measure. *Intelligence, 43*, 52–64. doi: 10.1016/j.intell.2014.01.004

Cortina, J. M., & Folger, R. G. (1998). When is it acceptable to accept a null hypothesis: No way, Jose? *Organizational Research Methods, 1*(3), 334–350. doi: 10.1177/109442819813004

Costa, P. T., & McCrae, R. R. (1992). Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-Factor Inventory (NEO-FFI) manual. Odessa, FL: Psychological Assessment Resources.

Cyert, R. M., & March, J. G. (1963). *The behavioral theory of the firm.* Englewood Cliffs, NJ: Prentice-Hall.

Dempsey, M. E. (2012). *Mission command white paper.* Washington, DC: Joint Chiefs of Staff.

Dragoni, L., Oh, I., Vankatwyk, P., & Tesluk, P. E. (2011). Developing executive leaders: The relative contribution of cognitive ability, personality, and the accumulation of work experience in predicting strategic thinking competency. *Personnel Psychology, 64*, 829–864. doi: 10.1111/j.1744-6570.2011.01229.x

DuBois, P. H. (1970). *A history of psychological testing*. Boston, MA: Allyn & Bacon.

Duhaime, I. M., Stimpert, L., & Chesley, J. A. (2012). *Strategic Thinking: Today's Business Imperative*. New York, NY: Routledge.

Dweck, C. (2006). *Mindset: The new psychology of success*. New York, NY: Random House, Inc.

Eden, C. (1990). Strategic thinking with computers. *Long Range Planning: International Journal of Strategic Management, 23*(6), 35–43. doi: 10.1016/0024-6301(90)90100-I

Eifler, B. S. (2012). *Developing strategic thinking leaders in the U.S. Army*. (Master's thesis). Carlisle, PA: U.S. Army War College.

Flanagan, J. C. (1954). The critical incident technique. *Psychological Bulletin, 41*, 237–358. doi: 10.1037/h0061470

Freedman, L. D. (2013). *Strategy: A history*. New York, NY: Oxford University Press.

Gerras, S. J. (2010). *Strategic Leadership Primer*. Carlisle Barracks, PA: United States Army War College.

Goldman, E. (2008). The power of work experience characteristics critical to developing expertise in strategic thinking. *Human Resource Development Quarterly, 19*(3). doi: 10.1002/hrdq.1237

Goldman, E. F., & Casey, A. J. (2010). Building a culture that encourages strategic thinking. *Journal of Leadership & Organizational Studies, 17*(2), 119–128. doi: 10.1177/1548051810369677

Goleman, D. (1995). *Emotional intelligence*. New York, NY: Bantam Books.

Graetz, F. (2002). Strategic thinking versus strategic planning: Towards understanding the complementarities. *Management Decision, 40*(5), 456–462. doi: 10.1108/00251740210430434

Gray, C. S. (2011). *The strategy bridge: Theory for practice.* Oxford, UK: Oxford University Press.

Gray, C. S. (2015). *The future of strategy*. Malden, MA: Polity Press.

Grier, R. A. (2012). Military cognitive readiness at the operational and strategic levels: A theoretical model for measurement development. *Journal of Cognitive Engineering and Decision Making, 6*(4), 358–392. doi: 10.1177/1555343412444606

Heracleous, L. (1998). Strategic thinking or strategic planning? *Long Range Planning: International Journal of Strategic Management, 31*(3), 481–487. doi: 10.1016/S0024-6301(98)80015-0

Huber, G. P. (2004). *The necessary nature of future firms: Attributes of survivors in a changing world.* Thousand Oaks, CA: Sage.

Huber, G. P. (2011). Organizations: Theory, design, and future. In S. Zedeck (Ed.) *APA Handbook of Industrial and Organizational Psychology, Vol. 1, Building and Developing the Organization* (pp. 117–160). Washington, DC: American Psychological Association.

Jacobs, T. O. (2002). *Strategic leadership: The competitive edge*. Washington, DC: National Defense University.

Kolb, D. (1984). *Experiential learning: Experience as the source of learning and development.* Upper Saddle River, NJ: Prentice Hall.

Krepinevich, A. F., & Watts, B. D. (2009). Regaining strategic competence. *Strategy for the Long Haul Series*. Washington, DC: Center for Strategic and Budgetary Assessments.

Landrum, R. E. (2011). Measuring dispositional humility: A first approximation. *Psychological Reports, 108*(1), 217–228. doi: 10.2466/02.07.09.PR0.108.1.217-228

Le, H., Oh, I., Robbins, S. B., Ilies, R., & Holland, E. (2011). Too much of a good thing: Curvilinear relationships between personality traits and job performance. *Journal of Applied Psychology, 96*(1), 113–133. doi: 10.1037/a0021016

Liedtka, J. (1998a) Strategic thinking: Can it be taught? *Long Range Planning: International Journal of Strategic Management, 31*(1), 120–129.

Liedtka, J. (1998b). Linking strategic thinking and strategic planning. *Strategy & Leadership, 26*(4), 30–35.

Lievens, F., & Motowidlo, S. J. (2016). Situational judgment tests: From measures of situational judgment to measures of general domain knowledge. *Industrial and Organizational Psychology, 9*(1), 3–22. doi: 10.1017/iop.2015.71

Lumpkin, G. T., & Dess, G. G. (1996). Clarifying the entrepreneurial orientation construct and linking it to performance. *Academy of Management Review, 21*(1), 135–172. doi: 10.2307/258632

Luthans, F. (2002). Positive organizational behavior: Developing and managing psychological strengths. *Academy of Management Executive, 16*, 57–72. doi: 10.2307/4165814

Maxfield, B. D. (2015). *Army demographics: FY15 Army profile*. Retrieved from http://www.armyg1.army.mil/hr/docs/demographics/BlueFinalFY15APR.pdf

Mayer, J. D., Salovey, P., & Caruso, D. R. (2000). Models of emotional intelligence. In R. Sternberg (Ed.), *Handbook of Intelligence* (pp. 396–420). Cambridge, UK: Cambridge University Press.

McCauley, D. (2012). Strategic thinking: Providing the competitive edge. *Small Wars Journal, 8*(2). Retrieved from http://smallwarsjournal.com/jrnl/art/strategic-thinking-providing-the-competitive-edge.

McDaniel, M. A., Hartman, N. S., Whetzel, D. L., & Grubb, W. L. (2007). Situational judgment tests, response instructions, and validity: A meta-analysis. *Personnel Psychology, 60*, 63–91. doi: 10.1111/j.1744-6570.2007.00065.x

McDaniel, M. A., Morgeson, F. P., Finnegan, E. B., Campion, M. A., & Braverman, E. P. (2001). Use of situational judgment tests to predict job performance: A clarification of the literature. *Journal of Applied Psychology, 80*, 730–740. doi: 10.1037//0021-9010.86.4.730

McDaniel, M. A., & Whetzel, D. L. (2007). Situational judgment tests. In D. L. Whetzel & G. R. Wheaton (Eds.), *Applied measurement*. (pp. 235–257). Mahwah, NJ: Erlbaum.

Mintzberg, H. (1994a). *The rise and fall of strategic planning*. New York, NY: The Free Press.

Mintzberg, H. (1994b). The fall and rise of strategic planning. *Harvard Business Review,* January–February, 107–114.

Morris, A. J., Brotheridge, C. M., & Urbanski, J. C. (2005). Bringing humility to leadership: Antecedents and consequences of leader humility. *Human Relations, 58*(10), 1323–1350. doi: 10.1177/0018726705059929

Moss, F.A. (1931). Preliminary report of a study of social intelligence and executive ability. *Public Personnel Studies,* 9, 2–9.

Motowidlo, S. J., Dunnette, M. D., & Carter, G. W. (1990). An alternative selection procedure: The low-fidelity simulation. *Journal of Applied Psychology, 75*, 640–647.

Motowidlo, S. J., Hooper, A. C., & Jackson, H. L. (2006). A theoretical basis for situational judgment tests. In J. A. Weekley & R. E. Ployhart (Eds.), *Situational judgment tests: Theory, measurement, and application* (pp. 57–81). Mahwah, NJ: Erlbaum.

Mumford, T. V., Van Iddekinge, C. H., Morgeson, F. P., & Campion, M. A. (2008). The team role test: Development and validation of a team role knowledge situational judgment test. *Journal of Applied Psychology, 93*(2), 250–267. doi: 10.1037/0021-9010.93.2.250

Office of the Deputy Assistant Secretary of Defense for Military Community and Family Policy (2014). *2014 demographics: Profile of the military community*. Retrieved from http://download.militaryonesource.mil/12038/MOS/Reports/2014-Demographics-Report.pdf.

Olson-Buchanan, J. B., & Drasgow, F. (2006). Multimedia situational judgment tests: The medium creates the message. In J. A. Weekley, R. E. Ployhart (Eds.) *Situational judgment tests: Theory, measurement, and application* (pp. 253-278). Mahwah, NJ: Lawrence Erlbaum Associates.

Oreg, S. (2003). Resistance to change: Developing an individual differences measure. *Journal of Applied Psychology, 88*(4), 680–693. doi: 10.1037/0021-9010.88.4.680

Payne, S. C., Youngcourt, S. S., & Beaubien, J. M. (2007). A meta-analytic investigation of the goal orientation nomological net. *Journal of Applied Psychology, 92*(1), 128–150. doi: 10.1037/0021-9010.92.1.128

Pelc, A., & Pelc, K. J. (2009). Same game, new tricks: What makes a good strategy in the prisoner's dilemma? *Journal of Conflict Resolution, 53*(5), 774–793. doi: 10.1177/0022002709339045

Pisapia, J. R., Reyes-Guerra, D., & Coukos-Semmel, E. (2005). Developing the leader's strategic mindset: Establishing the measures. *Leadership Review, 5*, 41–68.

Ployhart, R. E., & Holtz, B. C. (2008). The diversity-validity dilemma: Strategies for reducing racioethnic and sex subgroup differences and adverse impact in selection. *Personnel Psychology, 61*, 153–172. doi: 10.1111/j.1744-6570.2008.00109.x

Ployhart, R. E., & MacKenzie, W. I. (2011). Situational judgment tests: A critical review and agenda for the future. In S. Zedeck (Ed.) *APA Handbook of Industrial and Organizational Psychology, Vol. 2 Selecting and Developing Members for the Organization* (pp. 237–252). Washington, DC: American Psychological Association.

Ployhart, R. E., & Ryan, A. M. (2000). *Integrating personality tests with situational judgment tests for the prediction of customer service performance.* Paper presented at the 15th annual conference of the Society for Industrial and Organizational Psychology, New Orleans, LA.

Porter, M. E. (1987, May). Corporate strategy: The state of strategic thinking. *The Economist, 303*(7499), 17–24.

Porter, M. E. (1996). What is strategy? *Harvard Business Review, 74*(6), 61–78.

Porter, M. E. (2008). The five competitive forces that shape strategy. *Harvard Business Review, 86*(1).

Porterfield, R. W. (2001). *The development, psychometric evaluation, and validation of a situational judgment inventory for security officers.* (Doctoral Dissertation). Retrieved from ProQuest Dissertations & Theses Global. (3035702).

Reynolds, D. H., Sydell, E. J., Scott, D. R., & Winter, J. L. (2000, April). *Factors Affecting Situational Judgment Test Characteristics*. Paper presented at the 15th annual conference of the Society for Industrial and Organizational Psychology, New Orleans, LA.

Sacco, J. M., Schmidt, D. B., & Rogg, K. L. (2000, April). *Using readability statistics and reading comprehension scores to predict situational judgment test performance, black-white differences, and validity*. Paper presented at the 15th annual conference of the Society for Industrial and Organizational Psychology, New Orleans, LA.

Sackett, A. S., Karrasch, A. I., Weyhrauch, W. S., & Goldman, E. F. (2016). Enhancing the strategic capability of the Army: An investigation of strategic thinking tasks, skills, and development (Research Report 1995). Fort Belvoir, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. AD1006147.

Salmoni, B., Hart, J., McPherson, R., & Winn, K. (2010). Growing strategic leaders for future conflict. *Parameters, 40*, 72–88.

Salovey, P., Bedell, B. T., Detweiler, J. B., & Mayer, J. D. (2000). Current directions in emotional intelligence research. In M. Lewis & J. M. Haviland-Jones (Eds.), *Handbook of Emotions* (pp. 504–520). New York, NY: Guildford Press.

Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin, 124*(2), 262–274.

Schmitt, N. (1996). Uses and abuses of coefficient alpha. *Psychological Assessment,* 8, 350-353. doi:10.1037/1040-3590.8.4.350.

Schmitt, N., & Chan, D. (2006). *Situational judgment tests: Method or construct?* In J. A. Weekley & R. E. Ployhart (Eds.), *Situational judgment tests: Theory, measurement, and application* (pp. 135–155). Mahwah, NJ: Erlbaum.

Schwenck, C., Bjorklund, D. F., & Schneider, W. (2009). Developmental and individual differences in young children's use and maintenance of a selective memory strategy. *Developmental Psychology, 45*(4), 1034-1050. doi: 10.1037/a0015597.

Senge, P. M. (1990). *The fifth discipline: The art and practice of the learning organization.* New York, NY: Random House.

Sijtsma, K. (2009). On the use, the misuse, and the very limited usefulness of Cronbach's alpha. *Psychometrika, 74*(1), 107-120. doi: 10.1007/s11336-008-9101-0.

Sleesman, D. J., Conlon, D. E., McNamara, G., & Miles, J. E. (2012). Cleaning up the big muddy: A meta-analytic review of the determinants of escalation of commitment. *Academy of Management Journal, 55*(3), 541–562. doi:10.5465/amj.2010.0696

Smith, M. R. (2011). *Initial development and validation of the entrepreneurial orientation profile inventory (EOPI).* (Doctoral Dissertation). Retrieved from ProQuest Dissertations & Theses Global. (3458439).

Socin, R. J. (2008). *Measuring openness to experience in the work environment.* (Doctoral Dissertation). Retrieved from ProQuest Information and Learning. (AAI3298424).

Staw, B. M. (1976). Knee deep in the big muddy: A study of escalating commitment to a chosen course of action. *Organizational Behavior and Human Performance, 16,* 27–44.

Stieglitz, N., Knudsen, T., & Becker, M. C. (2015). Adaption and inertia in dynamic environments. *Strategic Management Journal*. Advance Online Publication. doi: 10.1002/smj.2433

Story, J. S. P., & Barbuto, J. E. (2011). Global mindset: A construct clarification and framework. *Journal of Leadership and Organizational Studies, 18*(3), 377–384. doi: 10.1177/1548051811404421

Sugerman, J., Scullard, M., & Wilhelm, E. (2011). *The 8 dimensions of leadership: DiSC strategies for becoming a better leader*. San Francisco, CA: Berrett-Koehler.

Torbert, B. (2004). *Action inquiry: The secret of timely and transforming leadership.* San Francisco, CA: Jossey-Bass.

U.S. Army Training and Doctrine Command (2009). *A Leader Development Strategy for a 21st Century Army*. Retrieved from http://www.cgsc.edu/alds/#imp1.

U.S. Department of Defense (2017). *Department of Defense Dictionary of Military and Associated Terms* (Joint Publication 1-02). Washington, DC: Author.

Van Rooy, D. L., & Viswesvaran, C. (2004). Emotional intelligence: A meta-analytic investigation of predictive validity and nomological net. *Journal of Vocational Behavior, 65*(1), 71–95.

Waters, D. E. (2011). Understanding strategic thinking and developing strategic thinkers. *Joint Forces Quarterly, 63*, 113–119.

Weekley, J. A., & Ployhart, R. E. (2006). An introduction to situational judgment tests. In J. A. Weekley and R. E. Ployhart (Eds.), *Situational judgment tests: Theory, measurement, and application* (pp. 1-10). Mahwah, NJ: Erlbaum.

Weekley, J. A., Ployhart, R. E., & Holtz, B. C. (2006). On the development of situational judgment tests: Issues in item development, scaling, and scoring. In J. A. Weekley & R. E. Ployhart (Eds.), *Situational judgment tests: Theory, measurement, and application* (pp. 135–155). Mahwah, NJ: Erlbaum.

Wheatley, W. J., Anthony, W. P., & Maddox, E. N. (1991). Selecting and training strategic planners with imagination and creativity. *The Journal of Creative Behavior, 25*(1), 52–60.

Whetzel, D. L., & McDaniel, M. A. (2009). Situational judgment tests: An overview of current research. *Human Resource Management Review, 19*, 188–202. doi: 10.1016/j.hrmr.2009.03.007

Wolters, H. M. K., Grome, A. P., & Hinds, R. M. (Eds.). (2013). *Exploring strategic thinking: Insights to assess, develop, and retain Army strategic thinkers* (Research Product 2013-1). Fort Belvoir, VA: U.S. Army Research Institute for the Behavioral and Social Sciences. ADA577290.

Yarger, H. (2008). *Strategy and the national security professional*. Westport, CT: Praeger Security International.

Yorks, L., & Nicolaides, A. (2012). A conceptual model for developing mindsets for strategic insight under conditions of complexity and high uncertainty. *Human Resource Development Review, 11*(2), 182–202. doi: 10.1177/1534484312439055

Young, R. E. (2004). *The creation of an emotional intelligence situational judgment test with the ability-based model of emotional intelligence as the operative construct*. (Doctoral Dissertation). Retrieved from ProQuest Dissertations & Theses Global. (3153979).

Zabriskie, N. B., & Huellmantel, A. B. (1991). Developing strategic thinking in senior management. *Long Range Planning: International Journal of Strategic Management, 24*(6), 25–32.

Zeidner, M., Matthews, G., & Roberts, R. D. (2004). Emotional intelligence in the workplace: A critical review. *International Review of Applied Psychology, 53*(3), 371–399.

Appendix A

Developing a Situational Judgment Test of Strategic Thinking Mindset

This research makes three primary contributions. First, this project contributes substantially to conceptual work in understanding strategic thinking by integrating theories of strategic thought and identifying the non-cognitive aspects of strategic thinking, covering the disciplines of psychology, management, and military doctrine and theory. Second, in developing an instrument capable of assessing the strategic thinking mindset, this research contributes to the selection, promotion, development, and mentoring of young officers in the Army. Given the extensive work required to produce SJT items and collect sufficient response data to assess item and scale functioning, this is an important contribution. Finally, the conceptualization, development, and outcomes of this effort may serve as a guide to other researchers in their own efforts to construct similar tests of strategic thinking mindset in other jobs and industries.

Developing a Measure

In order to supplement the literature on strategic thinking, specifically in regard to the mindset, as well as to provide the Army with a more context-relevant assessment product, this research was designed to develop a Strategic Thinking Mindset Test (STMT) using a situational judgment test (SJT) format. SJTs are characterized by a situation and response-choice format in which realistic situations from a designated job or work role are presented. Respondents are instructed to select a response option that represents the best/worst (or most/least likely) way to address the situation.

SJTs have grown in popularity among industrial and organizational psychologists (Ployhart & MacKenzie, 2011), particularly in selection and assessment contexts, thanks to a few particular strengths of SJTs. First, they tend to result in strong criterion-related validity for job performance (McDaniel, Hartman, Whetzel, & Grubb, 2007), although, the magnitude of this evidence is tied to the constructs measured. Additionally, there is evidence that SJTs result in smaller subgroup differences than traditional cognitive predictors (Ployhart & Holtz, 2008), resulting in reduced adverse impact.

Additionally, SJTs typically demonstrate high face validity (Ployhart & MacKenzie, 2011). By incorporating situations relevant to the work of the respondent, rather than generic interpersonal situations, SJTs stand a better chance of appearing valid and relevant to respondents and other stakeholders, compared to non-work context or context-free assessments common in psychological research. For example, a generic personality inventory might inquire about how one acts at social gatherings, an item that may raise concerns or appear to lack relevance to a respondent expecting to see items relevant to the workplace. While from a purely measurement perspective face validity has no real value, in organizational contexts, pragmatic concerns such as acceptability are important. The strategic thinking mindset assessment would, ideally, be used to guide self-reflection, self-development, and mentorship among military officers. However, without substantial face validity, as well as supportive construct validity evidence, the assessment would be less effective in this regard.

The first instruments designed to measure situational judgment appeared in the late 19th and early 20th centuries (DuBois, 1970; Moss, 1931). Research on SJT methods picked up after this form of testing was re-conceptualized as low-fidelity work simulation by Motowidlo, Dunnette, and Carter (1990). At first, SJTs were developed around critical job tasks identified from a job analysis, rather than a specific construct. Scientific investigations of the utility of job performance predictors revealed SJTs to be fairly strong predictors of performance (Schmidt & Hunter, 1998). More attention has since been paid to the idea that SJTs reflect a method of testing constructs, rather than a construct itself (Arthur & Villado, 2008; Schmitt & Chan, 2006). Therefore, an SJT's ability to predict performance or any other criterion depends on the construct it is designed to measure and that construct's relationship with the criterion.

Along that line, the use of SJTs has expanded into trait-based testing, with instruments being developed around a specific construct, other than the vaguely-defined situational judgment (Motowidlo, Hooper, & Jackson, 2006). Ployhart and MacKenzie (2011) note that a consistent method for developing SJTs around a single homogenous construct without variance related to cognitive ability or personality has yet to be developed. Therefore, SJTs typically correlate to some degree either with measures of cognitive ability or personality. However, there is evidence that choices in SJT development have a predictable effect on whether SJT scores correlate more with cognitive ability or personality (McDaniel et al., 2007). For example, response instructions (e.g., what *would* you do vs. what *should* you do) can cause an SJT to correlate more with personality and cognitive ability, respectively.

The SJT format enables the instrument to more directly assess the respondent's mindset by presenting a situation designed to evoke elements of the mindset. Relying on the critical incident technique to generate the scenarios ensures that the test benefits from a fidelity and experience of realism that would be difficult to replicate without the involvement of actual Army personnel. Their input on the types of scenarios that evoke the characteristics, the language that would be used to describe those scenarios, and the realistic range of response options are all compelling reasons to use a critical incident technique and an Army sample. By presenting response alternatives reflecting levels of flexibility, inclusiveness, and humility within realistic Army officer responses, the SJT data provide information about a respondent that is more context-bound than a behavioral self-report measure.

╕ Choices in Situational Judgment Test Development

There are a variety of alternatives available when developing an SJT, each of which affects the way the SJT functions and the constructs it is likely to measure. There are five main issues to be dealt with in developing an SJT: (a) item stem content, (b) response option content, (c) response instructions, (d) response keying, and (e) scoring methods (McDaniel & Whetzel, 2007; Weekley, Ployhart, & Holtz, 2006). Discussion of each of these issues as it relates to the current research follows.

Item Stem Content

The situation presented to the respondent is referred to as the item stem. The content of the stem is typically generated either by the researcher, based on theory, or by subject matter experts (SMEs) through the critical incident technique (Flanagan, 1954). A hybrid approach is

used in this research, by collecting incidents related to a specific context, which were then used to develop items for each mindset characteristic based on theory.

Stem complexity (e.g., detail, length, reading level) may also hinder or enhance criterion-related validity (McDaniel, Morgeson, Finnegan, Campion, & Braverman, 2001; Reynolds, Sydell, Scott, & Winter, 2000) and subgroup differences (Sacco, Schmidt, & Rogg, 2000). A strength of the critical incident approach used in this study, wherein the target population writes the initial incident description, is that the reading level for all scenarios begins at the level of the target population.

Stem fidelity (i.e., the extent to which the focal task is fully recreated) is another issue. Multimedia-based SJTs have taken advantage of this to eliminate reading requirements and enhance validity and respondent reactions (Olson-Buchanan & Drasgow, 2006). While multimedia SJTs have their advantages, development and implementation costs are high.

Finally, stem content can vary from interpersonal, judgment-based situations to more factual, procedural knowledge formats. This test is decidedly judgment-based, seeking to access the way the respondent deals with ambiguous situations, in regard to flexibility, inclusiveness, and humility. The development of item stem content is discussed further in the Stage 1 method and results section.

Response Option Content and Instructions

Paired with each item stem is a set of response options. The same issues of source, complexity, and fidelity (discussed in regard to item stem) also apply to response options. As with the item stem, the response options are construct-oriented, meaning each item focuses on either flexibility, inclusiveness, or humility. Following the example of Ployhart and Ryan (2000), a behavioral continuum approach is employed in which the response options reflect varying levels of the focal construct. The development of response option content is discussed further in the Stage 2 method and results section.

In regard to the instructions given for the SJT (i.e., framing of the question), there are typically two (or four) alternatives: Which option is your *most/least likely* response and which option is *the best/worst* (or *most/least effective*) response? SJTs based on likelihood instructions tend to correlate more with personality, while best/worst instructions tend to correlate more with cognitive ability (McDaniel et al., 2007). The mindset more closely resembles a personality trait than a cognitive ability; therefore, the test uses likelihood instructions.[2]

Response Keying

Once a set of items has been developed with stems and response options, a decision must be made about how to key each response in relation to the measured construct. In other words, each response must be identified as correct/incorrect or as an expression of some particular level of the construct. The keying scheme can be (a) empirically derived (e.g., based on each option's correlation with a criterion), (b) based on SME judgment, (c) rationally derived from theory, or

[2] Data were collected using best/worst instructions; the results showed no meaningful difference in correlation with cognitive ability.

(d) based on a hybrid of these approaches (Bergman et al., 2006; McDaniel & Whetzel, 2007). Empirical keying is ideal for situations with a clear performance criterion, which does not apply to this case. Purely following theory without any expert input would lack fidelity to real-world response tendencies. Therefore, a hybrid approach relying primarily on SME ratings, though informed by theory-based rationality, was used for this instrument. The development of the response keying system is described in the Stage 3 method and results section.

Scoring Methods

Finally, there are options related to how the SJT will be scored overall, translating the key into an item score. A forced-choice strategy involves a dichotomization of correct and incorrect responses, wherein a point is scored if a correct item response is endorsed. An expansion on this scoring approach is to tie a negative score to incorrect response options. As demonstrated by Motowidlo et al. (1990), the variance in a forced-choice item can be enhanced by scoring an item on a scale of −2 (best answer identified as worst and vice versa) to +2 (best and worst answers both correctly identified). Rather than a forced-choice strategy, others have adopted a continuous scale approach (e.g., Mumford, Van Iddekinge, Morgeson, & Campion, 2008), wherein respondents rate each response option on a scale of best/worst or most/least likely. The continuous rating method, however, complicates data analysis by shifting the level of analysis down, such that each stem becomes a testlet with items nested within it.

For this instrument, scoring was done with a forced-choice approach similar to the example provided by Motowidlo et al. (1990) to simplify scoring and reduce test length and respondent fatigue. The scoring system is described further in the Stage 4 method and results section.

Appendix B

Instructions for Reporting Critical Incidents

What do you mean by critical incident?

- A **specific example** of work, from your experience, representing some aspect of your job (in this case, a novel, complex problem
- They are ***not* descriptions of general types of problems** you experienced.
- There can be a **short or extended time frame**. Problems can be contained in a short period, or throughout a deployment.

What kind of critical incident do you want?

- Incidents of **dealing with highly complex problems requiring a novel solution**. In other words, unique problems that weren't specifically covered during training and don't have established solutions.
- If you did not experience such a problem, try to **think of one you closely observed** a peer dealing with.

What should I include?

- The **mindset** of an officer during decision-making in complex, novel problem situations. So, the **way you assessed the problem** is as relevant as **what you did to solve it**. Describe actions taken *and* the process for understanding the problem and possible solutions.
- Can be an example of **good, poor, or mediocre performance**. We want stories reflecting a range of performance. Complex, novel problems are rarely resolved perfectly; failures and successes are expected.
- We are **mostly concerned with the situations**. Focus the forces acting on the situation and how each impacted you.
- Courses of action that were considered, but not adopted, are welcome.

How will these be used?

- We will use the incidents you provide to **build problem scenarios** for officer development. Expert feedback and theory will help us **develop various courses of action**. These scenarios will help us compare and assess the way Officers respond to the types of scenarios presented.

Strategic Thinking Competencies

In addition to describing the behaviors you engaged in, you will be asked to indicate the types of thinking that were required, according to the categories defined below.

Critical Thinking: Reflecting on a variety of positions on an issue and evaluating each in an unbiased way. This type of thinking requires seeking out evidence to support a variety of points of view, clarifying assumptions, and making logical inferences.

Creative Thinking: Generating new insights into an issue. This type of thinking requires establishing new connections between previously unassociated concepts or applying them in new ways or in new contexts.

Systems Thinking: Perceiving how a set of elements interact to form a whole. This type of thinking requires considering how contextual variables influence the cause-and-effect relationships within an environment.

Thinking in Time: Considering the influence of the past and the desired future on the current status of an issue. This type of thinking requires examining how a current issue developed, how the pattern is likely to unfold in the future, and how actions could shape a desired future state.

Ethical Thinking: Considering the value, moral, and ethical factors in a situation which determine what one must do, what one must *not* do, and what is acceptable to do. This type of thinking requires knowledge of the values, morals, and ethics of yourself and those you represent.

Tips for writing good incidents

Describe an event, don't give advice. We need examples from the work of Army officers of complex, novel problems. Try to avoid speaking in general terms about how to solve such problems.

Write events in the third person (e.g., using "he" or "she" instead of "I" and "we"). Even if you are writing about things you did or things that happened to you, please write about them as things you observed.

Do not include specific identifying information, such as a person's name, a unit name, or a specific time and place (e.g. FOB Bravo, August 2010). Instead, use terms such as "the interpreter" or "MAJ X." To the extent that a detail is important (e.g. took place during a local election) include only enough to explain the context.

Take your time. It may take up to an hour to recall and report a relevant incident with sufficient detail. One effective incident is better than two that miss the point.

Include what you considered doing as well as what you did do.

DO NOT report incidents in which you participated in illegal activity. These may be subject to the Uniform Code of Military Justice and we may be required to report these events to the appropriate authorities.

Examples

On the following pages, there are two example incidents, to help you understand what we're looking for. The first is an example of an ideal style, length, and detail. The second is a real critical incident collected from a similar project we've conducted, but that is somewhat lacking in detail. We hope you can aim for something between these examples.

Example Incident 1

The following incident is a fictional depiction of an incident similar to what we are seeking. This is an adaptation of a scenario described in "The Strategic Corporal: Leadership in the Three Block War" by Marine Gen. Charles Krulak, in *Marines Magazine*, January 1999.

Check the box for each type of thinking that was required: ☐ Critical Thinking X Creative Thinking ☐ Systems Thinking X Thinking in Time X Ethical Thinking	
Situation (Refer to the instructions packet for clarification on what to include here)	CPL H was a squad leader in a Regional Multi-National Force (RMNF) providing security for a food distribution point (FDP). Food and medical supplies from the FDP had a positive impact on the community, with daily death tolls decreasing. A supply convoy brought news that members of a local faction, led by Warlord N, were gathering near the boundary of a rival faction's territory. Warlord N criticized the presence of the RMNF, but so far hadn't targeted Americans. However, as starvation became less of a concern, there was fear that political tensions would erupt in violence. LT F ordered CPL H's squad to man a roadblock at Checkpoint (CP) Charlie. Barricades were moved into place to secure the street. CPL S established an observation post on a nearby rooftop. A large crowd gathered waiting to pass through the checkpoint. CPL S reported that the crowd included many visibly upset young adult males. Meanwhile, he could see the vehicles of Warlord N's gang gathered at the boundary. CPL H learned that Warlord N's rival, Warlord M, was moving directly toward the CP and a likely collision with Warlord N, with the squad squarely in the middle. LT F directed CPL H to extend the road block and started moving another squad to help reinforce the checkpoint. The tension grew as the crowd became upset by the delay. The young men chanted anti-U.S. slogans and began to throw rocks. CPL H felt the situation slipping out of control and decided to close the road completely. The crowd erupted in protest and pressed forward.

	A helicopter nearby was hit by ground fire and crashed several blocks away. CPL S observed the crash and saw survivors. In the distance, he could see Warlord N's men rushing across the boundary. CPL S urgently requested permission to move to assist the crash survivors. Warlord M's armed men arrived at the checkpoint followed closely by local media. They forced their way up to the barricade and the crowd began pelting the squad with rocks. CPL H, who knew the fate of his men and perhaps the humanitarian mission, hung in the balance.
Problem-solving approach (Refer to the Instructions Packet for clarification on what to include here)	CPL H reviewed what he knew. He was certain that LT F and 2nd Squad would arrive within minutes. He knew that the crash site was in the adjacent unit's sector and that checkpoints existed along Warlord N's route to the site. He knew that exchange of gunfire with Warlord M's men would likely lead to civilian casualties and jeopardize the success of the humanitarian mission. Then, he considered what he didn't know. He wasn't certain of either warlord's intentions, nor of the likelihood of a successful rescue attempt. CPL S was directed to maintain his position and monitor Warlord N's progress and the status of the crash survivors. CPL H contacted the adjacent RMNF unit and learned that they had already dispatched medical personnel to the crash site. He informed them of Warlord N's movement, prompting them to reinforce the appropriate checkpoints.
Outcome (Refer to the Instructions Packet for clarification on what to include here)	LT F arrived with the additional squad, along with a neighborhood leader who had previously acted as an interpreter and mediator, recognized and respected in the community. Warlord M's men withdrew. The mediator addressed the crowd. The situation was diffused: Warlord M's men departed, the crowd calmed, and personnel reached the crash site.

Circle the number that best reflects how effectively the problem was dealt with

1	2	3	4	5
Very Ineffective/ Detrimental				Highly Effective/ Successful

Example Incident 2

The following is a real critical incident collected during a project investigating the work of combat advisors.

Check the box for each type of thinking that was required: X Critical Thinking ☐ Creative Thinking ☐ Systems Thinking ☐ Thinking in Time ☐ Ethical Thinking	
Situation (Refer to the instructions packet for clarification on what to include here)	There was an incident where CF fired on 4 IA who were on leave in civilian attire. US claimed these individuals were planting an IED however there was no evidence. 2 of the 4 were killed and the other two detained. Although our BN was not directly involved, our inability to satisfactory answer questions and relay decisions hurt our relationship with the IA BN officers. The CPT detained had been with the IA for four years. They passed through 11 checkpoints and had access to DIV HQs yet were accused of placing an IED in the middle of nowhere.
Problem-solving approach (Refer to the instructions packet for clarification on what to include here)	Advisor coordinated through CF Intel channels to ascertain detainee and incident information. CF held information in 15-6 format and would not release through intel channels. IA is certain this incident is a cover up by the US Army and these detainees will be held indefinitely.
Outcome (Refer to the instructions packet for clarification on what to include here)	CF refuses any wrong doing. 2 dead IA and 2 detained IA. Information provided from advisor to IA built a better relationship but lack of resolution damaged Iraqi/US relations. Advisor frustrated with no sense of accomplishment yet situation was resolved as far as the advisor could take it. This incident is like the entire deployment. Many tasks would not be considered complete by US standards. This mission is frustrating on the best day yet completed to the best ability of all advisors involved.
Circle the number that best reflects how effectively the problem was dealt with 1 2 3 4 5 Very Ineffective/ Detrimental ... Highly Effective/ Successful	

Incident 1

Using the space provided below, describe a problem-solving incident that you witnessed, either as an active participant or close observer. As described in the instructions packet, the incident should reflect a problem of heightened complexity and novelty. Pay special attention to detailing the nature of the situation and the various aspects contributing to the complexity.

Check the box for each type of thinking that was required: ☐ Critical Thinking ☐ Creative Thinking ☐ Systems Thinking ☐ Thinking in Time ☐ Ethical Thinking	
Situation (Refer to the instructions packet for clarification on what to include here)	
Problem-solving approach (Refer to the instructions packet for clarification on what to include here)	
Outcome (Refer to the instructions packet for clarification on what to include here)	
Circle the number that best reflects how effectively the problem was dealt with 1 2 3 4 5 Very Ineffective/ Detrimental ... Highly Effective/ Successful	

Incident 2

Using the space provided below, describe a problem-solving incident that you witnessed, either as an active participant or close observer. As described in the instructions packet, the incident should reflect a problem of heightened complexity and novelty. Pay special attention to detailing the nature of the situation and the various aspects contributing to the complexity.

Check the box for each type of thinking that was required: ☐ Critical Thinking ☐ Creative Thinking ☐ Systems Thinking ☐ Thinking in Time ☐ Ethical Thinking	
Situation (Refer to the instructions packet for clarification on what to include here)	
Problem-solving approach (Refer to the instructions packet for clarification on what to include here)	
Outcome (Refer to the instructions packet for clarification on what to include here)	
Circle the number that best reflects how effectively the problem was dealt with 1 2 3 4 5 Very Ineffective/ Detrimental ... Highly Effective/ Successful	

Demographics and Background

The information requested below will allow us to summarize the basic demographics and experiences of our sample.

1. Gender: _____

2. Rank (e.g., O-3): _____

3. Current functional branch (e.g. Civil Affairs, Infantry): _____

4. Current MOS/AOC (e.g., 12B): _____

Appendix C

Definitions

Flexibility

What it is:

- A mindset characterized by a willingness and tendency to adjust one's understanding, opinions, or approach when conditions change or new information is presented.

What it is NOT:

- Resisting necessary or optimal change.
    - Breaking from long-term strategy. "Bend, don't break" from a long-term strategy. *Bending* a strategy allows for adaptation and adjustment of the strategy and its implementation in the face of changing conditions, without abandoning the long-term effort and developing another strategy from scratch.

Why flexibility?

- Flexibility underlies the following themes of strategic thinking models:
    - *Adaptability and opportunism*: being willing and comfortable to deal with change, major or minor in nature, is crucial to being adaptive and taking advantage of opportunities.
    - *Action learning*: being inflexible might cause one to hesitate to look fairly at the results of an action, for fear that they might indicate changes are necessary.
    - *Creativity*: new ideas and associations will inevitably cause changes to one's understanding and processes in other areas, which requires a level of flexibility.

Inclusiveness

What it is:

- A mindset characterized by the welcoming of information and opinion from a broad range of sources (e.g. individuals, groups, disciplines, or other relevant parties).
- Maintaining an inclusive mindset allows one to value the holistic understanding that can come from examining an issue from many perspectives.

What it is NOT:

- Being hesitant to consider new or unusual sources of information.
- Fearing having too many voices involved in a discussion.
- Including everything; an inclusive mindset values the potential benefit of a broad perspective, while relying on other processes and judgment to filter information that does not contribute to understanding.

Why inclusiveness?

- Inclusiveness underlies the following themes of strategic thinking models:
    - *Breadth and inclusion*: engaging in a broad and inclusive information search requires first acknowledging and embracing the value of a broad information search.

- *Systems thinking and synthesis*: differs from traditional analytical processes in that system factors are gathered together for holistic understanding, rather than segmented into divisions which are handled separately. Therefore, systems thinking requires an inclusive mindset: a willingness to consider the breadth of a system all together.

Humility

What it is:

- A mindset characterized by comfort with admitting to being wrong or having an incomplete understanding of something.
- Tendency to check oneself; examining issues as if one's understanding is somehow wrong or incomplete.

What it is NOT:

- Reacting defensively to proposed changes or constructive feedback.
- Having a bias in favor of maintaining old beliefs or assumptions.

Why humility?

- Humility underlies the following themes of strategic thinking models:
  - *Self-awareness and self-control*: To critically evaluate one's own opinions and assumptions, a certain level of humility must be reached which allows for acceptance that one may be wrong.
  - *Criticality*: Once open to and comfortable with the position of reasonable self-doubt, behaviors and positions can be examined and, with effort, improved.

Scenario Review and Feedback Packet

Instructions

In this section, you will read a sequence of brief problem scenarios and provide feedback on their usability for the strategic thinking mindset assessment. You will also be asked to provide brief behavioral response options a Soldier might engage in to deal with that scenario. Additionally, we would like you to vary the level of flexibility, inclusiveness, or humility shown in the response options you provide.

Below is an example of a scenario with responses that vary on the element of inclusiveness. The A response is designed to reflect a low level of inclusiveness, while the B response reflects a high level of inclusiveness. Use these as a guide to the length and type of behavioral responses we are looking for.

EXAMPLE - You are a platoon leader and combat advisor assigned to mentor host nation police in an area of operations. You have two squads of coalition forces attached with you to a host nation police platoon. Another platoon nearby is delayed in arriving to their location, so your platoon is split in half and forced to cover two locations, reducing your combat effectiveness and placing your sergeant in charge at the second location. Both locations soon come under attack from positions deep in the surrounding mountains and valleys. The host national police at the other location are not responding to your Sergeant's instructions for dealing with the attack. He is quickly becoming frustrated and angry.

A. Travel to the second location to replace the Sergeant and give the instructions, as you've seen a tendency for the host nationals to have more respect for officers.
B. Consult with the interpreter to find out what may be causing the host nation forces to resist complying with the Sergeant's advice and what they think should be done instead.
C. Work on improving the defensibility of both positions while instructing the Sergeant to try other influence tactics and to control his emotions better when he speaks to the host nationals.
D. Instruct the Sergeant to have his squad execute the necessary actions for the host nationals until the other platoon arrives. Then report the incident to the host national police chief.

Scenarios

Instructions: Read the problem scenario provided in the box, then consider each question below. When appropriate, place an X in the box next to your desired response.

[Each scenario appeared in a text box like this, but for the sake of simplifying these appendices, the general form of the packet is presented only.]

Which of the three mindset characteristics is this scenario most appropriate for? In other words, is it easier to think of Soldier responses that vary on flexibility, inclusiveness, or humility?

□ Flexibility □ Inclusiveness □ Humility

In the space below, provide any general feedback you have on the scenario, how it might be fit to a mindset characteristic, made more relevant to a variety of Soldiers with different specialties, and any other additions or changes could be made to improve it.

(continues on other side)

Now, please consider what behaviors the Soldier in the above scenario might engage in to try to understand and resolve the problem.

In 1-2 sentences each, please describe up to ten different behavioral responses by varying the level of flexibility, inclusiveness, or humility shown, depending on which characteristic you believe can best be measured with this problem scenario. The responses don't have to be behaviors you think would definitely solve the problem, just actions that a typical Soldier might take.

1.

2.

3.

4.

5.

6.

7.

8.

9.

10.

Appendix D

Scenario and Response Rating Packet

Instructions

In this survey, you will be presented with real Company-grade problem scenarios. Associated with each is a set of response options that Army leaders have provided to us as realistic, viable ways to respond to the problem. Your task is to give us a rating of how effective each option would likely be. The second aspect of your task is to indicate how much each option indicates a mindset reflecting either flexibility, inclusiveness, or humility. Your ability to judge these scenarios and response options in the context of your experience and military training are crucial to developing a quality assessment.

For each scenario, you will be asked to rate how strongly you disagree or agree 1) that a response option would be effective, and 2) that the response option would reflect a mindset of flexibility/inclusiveness/humility. Please circle one number that best represents your level of agreement.

EXAMPLE - You are a platoon leader and combat advisor assigned to mentor host nation police in an area of operations. You have two squads of coalition forces attached with you to a host nation police platoon. Another platoon nearby is delayed in arriving to their location, so your platoon is split in half and forced to cover two locations, reducing your combat effectiveness and placing your sergeant in charge at the second location. Both locations soon come under attack from positions deep in the surrounding mountains and valleys. The host national police at the other location are not responding to your Sergeant's instructions for

1. Travel to the second location to replace the Sergeant and give the instructions, as you've seen a tendency for the host nationals to have more respect for officers.

(1) Strongly disagree	2 Somewhat disagree	3 Neither agree nor disagree	4 Somewhat agree	5 Strongly agree

Please ***do not*** use the space between numbers to indicate a more refined level of agreement, as shown below.

1. Consult with the interpreter to find out what may be causing the host nation forces to resist complying with the Sergeant's advice and what they think should be done instead.

1 Strongly disagree	2 Somewhat disagree	3 Neither agree nor disagree	4 Somewhat agree	5 Strongly agree

Scenarios

Please read the scenario and *all* of the response options *before* providing any ratings. Please rate each of these possible response options in terms of how strongly you disagree or agree that the response would be an ***effective way to address the problem***.

LT Nunez was an observer/controller (OC) tasked to oversee and provide feedback on a multinational, virtual exercise. Many tiers of systems participated, each with various sensors and weapon systems being used in the scenarios. During the scenarios, LT Nunez witnessed a glitch in the scenario, in which digital clutter appeared on the display screens whenever an incoming target was destroyed. The operators' responded by temporarily switching off their radar and turning it back on to remove the visual clutter.
LT Nunez saw this as a huge problem. The clutter and the switching off of the radar caused a lack of situational awareness on part of the operators, especially during intense parts of the scenario. LT Nunez brought his concerns to the attention of the scenario coordinator, but was met with resistance and an insistence that the lack of situational awareness was not a problem.

1. Gather evidence about the consequences of the glitch and bring it up during after-action review.

1 Strongly disagree	2 Somewhat disagree	3 Neither agree nor disagree	4 Somewhat agree	5 Strongly agree

2. Report the problem with the scenario coordinator to his/her superior.

1 Strongly disagree	2 Somewhat disagree	3 Neither agree nor disagree	4 Somewhat agree	5 Strongly agree

3. Talk to someone with more expertise in the systems being used and find out what it will take to fix the glitch.

1 Strongly disagree	2 Somewhat disagree	3 Neither agree nor disagree	4 Somewhat agree	5 Strongly agree

4. Accept that it is the scenario coordinator's job to oversee the running of the scenario and focus on how the participants overcome the lack of situational awareness.

1 Strongly disagree	2 Somewhat disagree	3 Neither agree nor disagree	4 Somewhat agree	5 Strongly agree

5. Focus on what the participants in the exercise are learning and don't hold problems associated with the glitch against them.

1 Strongly disagree	2 Somewhat disagree	3 Neither agree nor disagree	4 Somewhat agree	5 Strongly agree

6. Encourage the participants to complain that they don't feel they're getting a realistic exercise due to the glitch with the scenario systems.

1 Strongly disagree	2 Somewhat disagree	3 Neither agree nor disagree	4 Somewhat agree	5 Strongly agree

Review the scenario and response options as necessary and now please rate each of these possible options in terms of how strongly you agree or disagree that the response would ***reflect a mindset of humility***. Please refer back to the Definitions sheet you were provided to refresh yourself on what is meant by humility.

1. Gather evidence about the consequences of the glitch and bring it up during after-action review.

1 Strongly disagree	2 Somewhat disagree	3 Neither agree nor disagree	4 Somewhat agree	5 Strongly agree

2. Report the problem with the scenario coordinator to his/her superior.

1 Strongly disagree	2 Somewhat disagree	3 Neither agree nor disagree	4 Somewhat agree	5 Strongly agree

3. Talk to someone with more expertise in the systems being used and find out what it will take to fix the glitch.

1 Strongly disagree	2 Somewhat disagree	3 Neither agree nor disagree	4 Somewhat agree	5 Strongly agree

4. Accept that it is the scenario coordinator's job to oversee the running of the scenario and focus on how the participants overcome the lack of situational awareness.

1 Strongly disagree	2 Somewhat disagree	3 Neither agree nor disagree	4 Somewhat agree	5 Strongly agree

5. Focus on what the participants in the exercise are learning and don't hold problems associated with the glitch against them.

1 Strongly disagree	2 Somewhat disagree	3 Neither agree nor disagree	4 Somewhat agree	5 Strongly agree

6. Encourage the participants to complain that they don't feel they're getting a realistic exercise due to the glitch with the scenario systems.

1 Strongly disagree	2 Somewhat disagree	3 Neither agree nor disagree	4 Somewhat agree	5 Strongly agree