Harold F. O'Neil · Ray S. Perez
Eva L. Baker *Editors*

Teaching and Measuring Cognitive Readiness

Springer

Teaching and Measuring Cognitive Readiness

Harold F. O'Neil • Ray S. Perez • Eva L. Baker
Editors

Teaching and Measuring Cognitive Readiness

Springer

Editors
Harold F. O'Neil
University of Southern California
National Center for Research on Evaluation Standards, and Student Testing (CRESST)
Los Angeles, CA, USA

Ray S. Perez
Office of Naval Research
Arlington, VA, USA

Eva L. Baker
National Center for Research on Evaluation, Standards, and Student Testing (CRESST)/UCLA
Los Angeles, CA, USA

ISBN 978-1-4614-7578-1 ISBN 978-1-4614-7579-8 (eBook)
DOI 10.1007/978-1-4614-7579-8
Springer New York Heidelberg Dordrecht London

Library of Congress Control Number: 2013942797

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

In this edited book, we will focus on documenting recent theory and research on the teaching and assessment of cognitive readiness. What is cognitive readiness? Although there are many definitions of this construct, as indicated in the chapters in this book, we view cognitive readiness through a knowledge, skills, and attributes (KSA) lens (McLagan, 1997), that is, knowledge is domain specific, skills are either domain specific or domain independent, but attributes are relatively domain independent. Attributes are considered as widely applicable but hard to train. The term attribute is usually considered to be interchangeable with the term competency. However, Klieme, Hartig, and Rauch (2008) provide an interesting alternative view of competence reflecting mainly a European view.

Figure 1 provides a graphic representation of the constructs that compose cognitive readiness. This figure also provides the overall conceptual framework that drove the selection of authors to write book chapters by the editors. As may be seen in Fig. 1 the key constructs in our model are shown and consist at the top level as KSA. Knowledge includes domain-specific knowledge for developing cognitive readiness in specific domains as well as prerequisite skills. There are five cognitive readiness skills: adaptability, adaptive problem solving, communication, decision making, and situation awareness. There are four competencies, that is, adaptive expertise, creative thinking, metacognition, and teamwork.

Our definitions of these constructs are provided in Table 1. There are many different but closely related definitions of cognitive readiness that are used in this book. Each author was requested to be explicit regarding his or her conception of the constructs that compose cognitive readiness.

Chapters in this edited book vary from broad theoretical views to more narrow in-depth descriptions of specific subconstructs composing cognitive readiness. This book is organized into two major sections: theory/context and cognitive readiness applications.

The theory/context section (Chaps. 1–10) provides a rich description of cognitive readiness and its various definitions, models, and theories, as well as models for its teaching and assessment. The specific cognitive readiness constructs of adaptability, adaptive problem solving, situation awareness, and adaptive expertise are

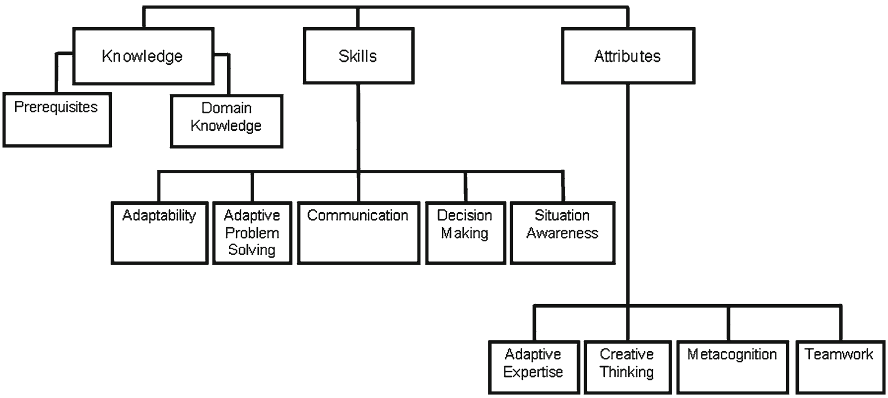


Fig. 1 Cognitive readiness model

Table 1 O'Neil's cognitive readiness skills and attributes

Skills and attributes	Definition
Adaptability	Adaptability is a functional change (cognitive, behavioral, and/or affective) in response to actual or correctly anticipated alterations in environmental contingencies (Banks, Bader, Fleming, Zaccaro, & Barber, 2001, p. 4)
Adaptive expertise	Adaptive expertise entails a deep understanding of the knowledge of a problem domain. Adaptive experts understand when and why particular knowledge is appropriate or not (Zaccaro & Banks, 2004; Ericsson, this volume)
Adaptive problem solving	Adaptive problem solving involves the ability to invent solutions to problems that the problem solver has not encountered before. In adaptive problem solving, problem solvers must adapt their existing knowledge to fit the requirements of a novel problem (Mayer, this volume). Adaptive problem solving has also been conceptualized by O'Neil (1999) as being composed of content understanding, problem solving strategies, and self-regulation
Communication	Communication is the timely and clear provision of information (Bowers, Braun, & Morgan, 1997) and the ability to know whom to contact, when to contact, and how to report (Hussain, Bowers, & Blasko-Drabik, this volume)
Creative thinking	Creative thinking is a predictor of creative accomplishment. Creative thinking is the ability to generate ideas and solutions that are novel, appropriate, and of high quality (Hong & Milgram, 2008)
Decision making	Decision making involves the use of situational awareness information about the current situation to help evaluate the utility of potential courses of action and then execute a course of action and judge its effectiveness. It involves the ability to follow appropriate protocols, follow orders, and take the initiative to complete a mission (Hussain et al., this volume)

(continued)

Table 1 (continued)

Skills and attributes	Definition
Metacognition	Metacognition is awareness of one's thinking and is composed of two components: planning and self-monitoring. Planning means that one must have a goal (either assigned or self-directed) and a plan to achieve the goal. Self-monitoring means one needs a self-checking mechanism to monitor goal achievement (O'Neil, 1999)
Situation awareness	Situation awareness involves being aware of what is happening around you, to understand how information, events, and your own actions will affect your goals and objectives, both now and in the near future. More formally, situation awareness can be defined as the perception of elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future (Endsley, 1995, p. 36)
Teamwork	Teamwork is a trait of the individual that predisposes the individual to act as a team member. There are six teamwork processes: (a) adaptability, (b) coordination, (c) decision making, (d) interpersonal, (e) leadership, and (f) communication (O'Neil, Wang, Lee, Mulkey, & Baker, 2003). A complementary definition is provided by Bowers and Cannon-Bowers in this volume. Their definition of teamwork includes knowledge of teamwork, leadership, mutual performance monitoring/back-up, communication, interpersonal skills, and positive teamwork attitudes

also presented. The section closes with chapters on twenty-first century skills and a cognitive readiness prerequisite skill.

The cognitive readiness applications section (Chaps. 11–17) provides both empirical and theoretical data of creative thinking, the use of analogies for instruction, and the use of simulations for teaching and assessment. This section concludes with two chapters related to software support and team training for cognitive readiness. In summary, this section's chapters are a synthesis of both empirical as well as theoretical views of specific cognitive readiness constructs.

The chapters also reflect the following issues: (1) a focus on a KSA view of cognitive readiness; (2) a focus on individual cognitive readiness KSA rather than team cognitive readiness (the exceptions are the chapters by Bowers and Cannon-Bowers, and Hussain et al.); (3) contexts in both schools and the workplace; (4) multiple approaches for assessment; (5) a focus on validity and cost; and (6) common low-stakes assessment purposes, i.e., diagnostic, program evaluation, or accountability.

Los Angeles, CA, USA Harold F. O'Neil
Arlington, VA, USA Ray S. Perez
Los Angeles, CA, USA Eva L. Baker

References

Banks, D., Bader, P. K., Fleming, P. J., Zaccaro, S. J., & Barber, H. F. (2001). *Leader adaptability: The role of work experiences and individual differences*. Paper presented at the 16th annual conference of the Society for Industrial and Organizational Psychology, San Diego, CA.

Bowers, C. A., Braun, C. C., & Morgan, B. B. (1997). Team workload: Its meaning and measurement. In M. T. Brannick, E. Salas, & C. Prince (Eds.), *Team performance assessment and measurement: Theory, methods, and applications* (pp. 85–108). Mahwah, NJ: Erlbaum.

Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors, 37*, 32–64.

Hong, E., & Milgram, R. M. (2008). *Preventing talent loss*. New York: Routledge.

Klieme, E., Hartig, J., & Rauch, D. (2008). The concept of competence in educational contexts. In J. Hartig, E. Klieme, & D. Leutner (Eds.), *Assessment of competencies in educational contexts* (pp. 3–22). Cambridge, MA: Hogrefe & Huber.

McLagan, P. A. (1997). Competencies: The next generation. *Training & Development, 51*(5), 40–47.

O'Neil, H. F., Jr. (1999). Perspectives on computer-based performance assessment of problem solving: Editor's introduction. *Computers in Human Behavior, 15*, 255–268.

O'Neil, H. F., Jr., Wang, S., Lee, C., Mulkey, J., & Baker, E. L. (2003). Assessment of teamwork skills via a teamwork questionnaire. In H. F. O'Neil, Jr., & R. S. Perez (Eds.), *Technology applications in education: A learning view* (pp. 283–303). Mahwah, NJ: Erlbaum.

Zaccaro, S. J., & Banks, D. (2004). Leader visioning and adaptability: Bridging the gap between research and practice on developing the ability to manage change. *Human Resource Management, 43*(4), 367.

Acknowledgements

Preparation of this edited book was supported in part by grants from the Office of Naval Research (ONR), Award Number N00014-08-1-0126 and Award Number N00014-09-C-0813. The work was supported by subcontracts from the University of California, Los Angeles (UCLA). The findings and opinions expressed in this book do not necessarily reflect the positions or policies of the Office of Naval Research. Special thanks to Ms. Joanne Michiuye, Mr. Rob Sitter, and Ms. Laquita Stewart for their excellent assistance in preparation of this manuscript.

Contents

Contributors

Eva L. Baker National Center for Research on Evaluation, Standards, and Student Testing (CRESST)/UCLA, Los Angeles, CA, USA

William L. Bewley National Center for Research on Evaluation, Standards, and Student Testing (CRESST)/UCLA, Los Angeles, CA, USA

Holly Blasko-Drabik Department of Psychology, University of Central Florida, Orlando, FL, USA

Cheryl A. Bolstad SA Technologies, Inc., Forest Hill, MD, USA

Clint Bowers Department of Psychology, University of Central Florida, Orlando, FL, USA

Klaus Breuer Johannes Gutenberg-Universität, Am Rodelberg Mainz (JGU), Germany

Hongwen Cai Guangdong University of Foreign Studies, Guangzhou, People's Republic of China

Janis Cannon-Bowers Institute for Simulation & Training, University of Central Florida, Orlando, FL, USA

Tiffani R. Chen Department of Psychology, George Mason University, Fairfax, VA, USA

Gregory K. W. K. Chung National Center for Research on Evaluation, Standards, and Student Testing (CRESST)/University of California, Los Angeles (UCLA), Los Angeles, CA, USA

Haydee M. Cuevas Embry-Riddle Aeronautical University, Orlando, FL, USA

Girlie C. Delacruz National Center for Research on Evaluation, Standards, and Student Testing (CRESST)/University of California, Los Angeles (UCLA), Los Angeles, CA, USA

Mica R. Endsley SA Technologies, Inc., Gold Canyon, AZ, USA

K. Anders Ericsson Department of Psychology, Florida State University, Tallahassee, FL, USA

Donna Escalante University of Southern California, Los Angeles, CA, USA

J.D. Fletcher Science and Technology Division, Institute for Defense Analyses, Alexandria, VA, USA

F Sutter Fox University of Southern California, Los Angeles, CA, USA

Joan L. Herman National Center for Research on Evaluation, Standards, and Student Testing (CRESST), UCLA Graduate School of Education and Information Studies, Los Angeles, CA, USA

Keith J. Holyoak Department of Psychology, University of California, Los Angeles, CA, USA

Eunsook Hong Department of Educational Psychology, University of Nevada, Las Vegas, NV, USA

Talib S. Hussain Raytheon BBN Technologies, Cambridge, MA, USA

Barbara Jones National Center for Research on Evaluation, Standards, and Student Testing (CRESST)/UCLA, Los Angeles, CA, USA

Joan (Yuan-Chung) Lang University of Southern California, Los Angeles, CA, USA

John J. Lee National Center for Research on Evaluation, Standards, and Student Testing (CRESST)/UCLA, Los Angeles, CA, USA

Michael D. Matthews Department of Behavioral Sciences and Leadership, U.S. Military Academy, West Point, NY, USA

Richard E. Mayer Department of Psychology, University of California, Santa Barbara, Santa Barbara, CA, USA

Allen Munro Center for Cognitive Technology, Rossier School of Education, University of Southern California, Redondo Beach, CA, USA

Harold F. O'Neil University of Southern California, National Center for Research on Evaluation Standards and Student Testing (CRESST), Los Angeles, CA, USA

Ray S. Perez Office of Naval Research, Arlington, VA, USA

Lindsey E. Richland Department of Comparative Human Development, University of Chicago, Chicago, IL, USA

Robert J. Sternberg Office of Academic Affairs, Oklahoma State University, Stillwater, OK, USA

Siegfried Streufert Professor Emeritus at Penn State College of Medicine, Hershey, PA, USA

Eric J. Weis Simon Center for the Professional Military Ethic, United States Military Academy, West Point, NY, USA

Alexander P. Wind State University of New York, Albany, NY

Stephen J. Zaccaro Department of Psychology, George Mason University, Fairfax, VA, USA

Part I
Theory/Context

Chapter 1
What Is Cognitive Readiness?

Harold F. O'Neil, Joan (Yuan-Chung) Lang, Ray S. Perez, Donna Escalante, and F Sutter Fox

1.1 The Concept and Context of Cognitive Readiness

In this book there are multiple views of cognitive readiness. The purpose of this chapter is to review the various definitions of cognitive readiness and also the definition of the constructs that compose cognitive readiness. The construct of cognitive readiness will also be briefly compared and contrasted with twenty-first century skills. Cognitive readiness denotes the mental preparation for effective changes in response to altered or unpredictable situations in this fast-changing world (Fletcher, 2004; Fletcher & Wind, this book).

Fletcher (2004) defined cognitive readiness as the mental preparation (including skills, knowledge, abilities, motivations, and personal dispositions) an individual needs to establish and sustain competent performance in a complex and unpredictable environment (e.g., modern military operations). Goldberg (2012), Fletcher (2009), and Laurence and Matthews (2012) provide excellent discussions of the characteristics of such a military environment.

In the Department of Defense (DoD), the term readiness denotes that achieving readiness will ensure that the war fighter is mentally prepared to be combat effective to accomplish the mission (Etter, Foster, & Steele, 2000; Spiering et al., 2012). The term readiness is also used in preparation for schooling, e.g., reading readiness

H.F. O'Neil (✉)
University of Southern California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST), Los Angeles, CA, USA
e-mail: honeil@usc.edu

J.(Y.-C.) Lang • D. Escalante • F S. Fox
University of Southern California, Los Angeles, CA, USA
e-mail: joanyclang@gmail.com; drayalaescalante@gmail.com; SutterFox@yahoo.com

R.S. Perez
Office of Naval Research, Arlington, VA, USA
e-mail: perezr@onr.navy.mil

H.F. O'Neil et al. (eds.), *Teaching and Measuring Cognitive Readiness*,
DOI 10.1007/978-1-4614-7579-8_1, © Springer Science+Business Media New York 2014

(see Herman's chapter in this book). A similar construct is twenty-first century skills (Griffin, McGaw, & Care, 2012; Mayrath, Clarke-Midura, & Robinson, 2012). Baker (in this book) provides a nuanced discussion of the similarities and differences between cognitive readiness and twenty-first century skills.

Some researchers look at cognitive readiness from a comprehensive perspective. For example, in the Fletcher and Wind chapter in this book, they refine Fletcher's (2004) conceptualization and define cognitive readiness to include adaptability, communication, creativity, critical thinking, decision making, metacognition, pattern recognition, problem solving, resilience, situational awareness, and teamwork and interpersonal skills. Bolstad, Cuevas, Babbitt, Semple, and Vestewig (2006) and Bolstad, Endsley, and Cuevas (this volume) also have a comprehensive view of 21 characteristics that define cognitive readiness, e.g., behavioral style, cognitive resources, cohesion, commonality of goals, communication, conflict management, decision making, emotion, fatigue, and flexibility.

O'Neil's model conceptualizes cognitive readiness into three major categories: (1) knowledge, (2) skills, and (3) attributes (KSA) (McLagan, 1997). In this framework, knowledge is domain specific, skills are either domain specific or domain independent, but attributes are relatively domain independent. Attributes are considered as widely applicable but harder to train (Jackson, Thoemmes, Jonkmann, Ludtke, & Trautwein, 2012). The term attribute is usually considered to be interchangeable with the term competency. However, Klieme, Hartig, and Rauch (2008) provide an alternative view of competence reflecting mainly a European view. Attribute is also often used interchangeably with the term attitude. We view attributes as more inclusive than the term attitude; thus, the "A" in our KSA model refers to attributes.

Building on these past research efforts, we revised Fletcher's model (2004; Fletcher & Wind, this volume) and developed O'Neil's Cognitive Readiness Model for the training and assessment of cognitive readiness. For our Cognitive Readiness Model, several components in Fletcher's model were dropped and new components were added. The component of pattern recognition was dropped as we considered it a basic psychological process with a high genetic nature and therefore difficult to modify. The components of interpersonal skills and resilience were also dropped as they were conceptualized by us to be more of an affective or feeling component and less of a cognitive-related attribute. We also dropped critical thinking as this construct subsumes many skills and attributes and thus would be difficult to define, assess, and teach (Abrami et al., 2008; West, Toplak, & Stanovich, 2008). Finally we added adaptive expertise.

We agree with Fletcher and Wind (in this volume) that the components should be relatively content- and context-free, be measureable, and be trainable. While Fletcher and Wind characterized cognitive readiness at multiple levels, i.e., a personal level, a team level, or an organizational level, our Cognitive Readiness Model focused on individuals rather than teams or organizations, e.g., "workers to be trained," "Navy personnel to be trained," or "college students to be instructed." Like Fletcher and Wind's requirement that cognitive readiness should be trainable, a constraint for our Cognitive Readiness Model was that the components of cognitive readiness would possess instructional sensitivity (see Baker, 2008, 2012 for her characterization of instructional sensitivity).

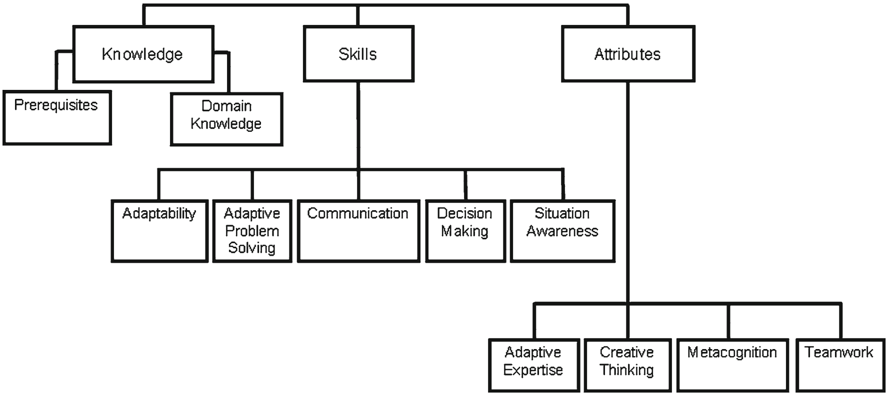


Fig. 1.1 Cognitive readiness model

Several training and assessment issues regarding facilitating transfer for cognitive readiness will be presented in this chapter. Many of the instructional strategies (e.g., train with worked examples, Shen & O'Neil, 2008; Sweller, Ayres, & Kalyuga, 2011) and measurement strategies (e.g., use of simulation in this chapter) that are necessary to train individuals have a firm empirical basis (e.g., worked examples), but others require more empirical validation (e.g., simulation). Overall, the training and assessment of cognitive readiness skills and attributes are in need of extensive research.

Our Cognitive Readiness Model provides a comprehensive framework for understanding, training, and assessing cognitive readiness. The key constructs in our model are shown in Fig. 1.1 and consist at the top level of the figure as knowledge, skills, and attributes. Knowledge includes domain-specific knowledge for developing cognitive readiness in specific content/process domains as well as prerequisite skills. There are five skills—adaptability, adaptive problem solving, communication, decision making, and situation awareness—and four attributes, i.e., adaptive expertise, creative thinking, metacognition, and teamwork. Table 1.1 provides definitions of these constructs. These definitions are important as they drive measurement.

Table 1.1 has been adapted from Morrison and Fletcher (2001) and provides our definitions for cognitive readiness skills and attributes.

1.1.1 Criticality of Cognitive Readiness

1.1.1.1 The Changing World

The United States is well into a transformation from an industrial-age economy to an information-age one. The information age is marked by the increased production, transmission, consumption of, and reliance on information. Social, cultural, and economic patterns reflect the decentralized, nonhierarchical flow of information

Table 1.1 O'Neil's cognitive readiness skills and attributes

Skills and attributes	Definition
Adaptability	Adaptability is a functional change (cognitive, behavioral, and/or affective) in response to actual or correctly anticipated alterations in environmental contingencies (Banks, Bader, Fleming, Zaccaro, & Barber, 2001, p. 4)
Adaptive expertise	Adaptive expertise entails a deep understanding of the knowledge of a problem domain. Adaptive experts understand when and why particular knowledge is appropriate or not (Zaccaro & Banks, 2004; Ericsson, this volume)
Adaptive problem solving	Adaptive problem solving involves the ability to invent solutions to problems that the problem solver has not encountered before. In adaptive problem solving, problem solvers must adapt their existing knowledge to fit the requirements of a novel problem (Mayer, this volume). Adaptive problem solving has also been conceptualized by O'Neil (1999) as being composed of content understanding, problem-solving strategies, and self-regulation
Communication	Communication is the timely and clear provision of information and the ability to know whom to contact, when to contact, and how to report (Hussain, Bowers, & Blasko-Drabik, this volume; Bowers, Braun, & Morgan, 1997)
Creative thinking	Creative thinking is a predictor of creative accomplishment. Creative thinking is the ability to generate ideas and solutions that are novel, appropriate, and of high quality (Hong & Milgram, 2008)
Decision making	Decision making involves the use of situation awareness information about the current situation to help evaluate the utility of potential courses of action and then execute a course of action and judge its effectiveness. It involves the ability to follow appropriate protocols, follow orders, and take the initiative to complete a mission (Hussain et al., this volume)
Metacognition	Metacognition is awareness of one's thinking and is composed of two components: planning and self-monitoring. Planning means that one must have a goal (either assigned or self-directed) and a plan to achieve the goal. Self-monitoring means one needs a self-checking mechanism to monitor goal achievement (O'Neil, 1999)
Situation awareness	Situation awareness involves being aware of what is happening around you, and understanding how information, events, and your own actions will affect your goals and objectives, both now and in the near future. More formally, situation awareness can be defined as the perception of elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future (Endsley, 1995, p. 36)
Teamwork	Teamwork is a trait of the individual that predisposes the individual to act as a team member. There are six teamwork processes: (a) adaptability, (b) coordination, (c) decision making, (d) interpersonal, (e) leadership, and (f) communication (O'Neil, Wang, Lee, Mulkey, & Baker, 2003). A complementary definition is provided by Bowers and Cannon-Bowers in this volume. Their definition of teamwork includes knowledge of teamwork, leadership, mutual performance monitoring/backup, communication, interpersonal skills, and positive teamwork attitudes

(e.g., Kozlowski, Gully, Salas, & Cannon-Bowers, 1996; OECD, 2012; Schneider, 1994). Schleicher (2006) wrote, "The world is indifferent to tradition and past reputations, unforgiving of frailty and ignorant of custom or practice. Success will go to those individuals and countries which are swift to adapt, slow to complain, and open to change" (Schleicher, 2006, p. 16).

Global competition, new production techniques, and rapid technological change have placed a premium on creativity and innovation (Baker, 2012; Mumford, Scott, Gaddis, & Strange, 2002). Organizations have found that success requires shifting from routine and highly scripted jobs to adapting the skills and abilities of people to evolving demands (Carnevale, Jayasundera, & Cheah, 2012). Perhaps the most valuable skill or attribute for today's workforce is one's cognitive readiness that allows transfer of learning from one system or scenario to another often without formal retraining.

1.1.1.2 The US Military as an Example

In the US Army, one of the most predictable characteristics of military operations is their unpredictability (Fletcher, 2004). The current military operating environment has become extremely complex. The complexity has increased for many reasons—the rapid advances and proliferation of technology, the dispersal and independent operation of military units, the intermingling and interdependence of military and civilian activities and personnel, and the presence of asymmetric threats. Faced with continual shifts in responsibilities, tasks, and missions (Fletcher, 2004), adaptive leaders learn to live with unpredictability and focus more on exploiting opportunities (Wong, 2004). In the Army, for example, a key factor in developing adaptive capacity in junior officers is the ability to actually lead and make decisions consistent with the commander's intent rather than merely to execute the orders of higher commands. An excellent description of this process is provided in TRADOC pamphlet 525-8-2 (2011).

1.2 O'Neil's Cognitive Readiness Model

1.2.1 The Cognitive Readiness Skills and Attributes

This section of the chapter will expand upon the definition provided in Table 1.1 and depicted in Fig. 1.1. The intent is to offer the reader a window into the recent literature regarding these constructs. As was indicated in Fig. 1.1, the five cognitive readiness skills are the following:

1. Adaptability
   Adaptability is an effective change in response to an altered situation (Mueller-Hanson, White, Dorsey, & Pulakos, 2005). Adaptability allows individuals to respond quickly and intelligently to constant change. It allows individuals to act

and then evaluate results instead of attempting to collect and analyze all the data before acting (Bennis, 2003; Tucker & Gunther, 2009). The idea of adaptive performance is multidimensional. Dimensions of adaptability include mental adaptability, interpersonal adaptability, physical adaptability, and cultural adaptability (Johnson, Friedland, Watson, & Surface, 2012; Pulakos, Arad, Donovan, & Plamondon, 2000). Since our Cognitive Readiness Model focuses on the cognitive component, other types of adaptability, e.g., intercultural competency (Johnson et al., 2012), will not be assessed.

2. Adaptive problem solving
   Problem solving is cognitive processing directed at transforming a given situation into a desired situation when no obvious method of solution is available to the problem solver; a problem exists when a problem solver has a goal but does not know how to reach it, so problem solving is a mental activity aimed at finding a solution to a problem (Baker & Mayer, 1999). In adaptive problem solving, problem solvers must adapt their existing knowledge to fit the requirements of a novel problem, including a novel problem situation that may change over time (Mayer, this volume). According to Mayer (2002), adaptive problem solving is one of the most significant skills whether in job settings or in schools, and, as a result, teaching and assessing problem solving have become one of the most significant educational objectives. Mayer (this volume) also points out based on Anderson et al. (2001) that five kinds of knowledge support cognitive readiness for adaptive problem solving—facts, concepts, procedures, strategies, and beliefs.

   O'Neil's Problem Solving Model (O'Neil, 1999) is based on Mayer's conceptualization for problem solving. Mayer's framework has been further refined by O'Neil (1999) into three components—content understanding, problem-solving strategies, and self-regulation. It is hypothesized that each of the three components can be trained and assessed independently. Content understanding indicates understanding of domain knowledge. The problem-solving strategies encompass both domain-specific and domain-independent problem-solving strategies. Metacognition is composed of planning and self-monitoring. This model suggests that to be successful problem solvers, individuals must know content, possess strategies to analyze and evaluate options (some of which apply specifically to the content in question), and be able to plan and monitor their progress toward the solution (Baker & O'Neil, 2003).

3. Communication
   Individual communication skills are associated with performance, especially in a teamwork environment. The quality and the amount of communication and even nonverbal communications are of crucial importance (Bowers & Cannon-Bowers, this volume). This particular cognitive readiness skill is not addressed in detail in this volume and will not be discussed further.

4. Decision making
   Decision making is the cognitive process leading to the selection of a course of action among variations. Situation awareness is the front end to a decision-making process. Effective decision making requires extensive domain-specific knowledge, such as mental models that describe causal relationships among events in the domain (Cohen et al., 2000, pp. 32–33). Different methodological

approaches have been used for decision making (Finucane, Mertz, Slovic, & Schmidt, 2005).

Normative models of decision making typically identify four fundamental skills: (1) belief assessment that involves judging the likelihood of outcomes (note in our model, situation awareness captures this skill), (2) value assessment that involves evaluating outcomes, (3) integration that involves combining beliefs and values in making decisions, and (4) metacognition that means knowing the extent of one's abilities (e.g., Finucane & Lees, 2005). These models judge the quality of a decision by its process rather than by its outcome, as it is assumed that a person who uses better decision processes will be more likely to experience good decision outcomes (e.g., Keren & Bruine de Bruin, 2003; Worthy, Gorlick, Pacheco, Schnyer, & Maddox, 2011).

An alternative view of decision making is proposed by Klein (2008). He demonstrated that experts respond to situations based on a stored repertoire of responses. When people recognize a situation as being similar to past situations they have encountered, they draw on the responses associated with such situations and respond almost automatically or intuitively. Such "naturalistic decision-making" processes are particularly effective in high-pressure and ambiguous situations, when time pressure prohibits a more structured, normative approach (Mueller-Hanson et al., 2005). This view of decision making relies heavily on prior knowledge and expertise.

5. Situation awareness
Situation awareness is generally defined as the ability to perceive and comprehend oneself in relation to relevant elements of the present environment and then accurately project different courses of action into the future (Endsley, 1988; see also Bolstad et al. in this volume). Endsley (1995) considered situation awareness to be a picture, "product," or mental model of the situation and not the cognitive processes underlying and supporting it per se.

Endsley's (1995) model of situation awareness contains three major components: perception, comprehension, and projection. In order to achieve situation awareness, the individual must first perceive the relevant elements from the environment. Comprehension is the cognitive phase responsible for cementing together the disparate elements (building blocks) acquired from the environment with prior knowledge in order to comprehend the present situation. Projection is defined by the ability to predict the future system states and plan ahead based on the comprehension and understanding of the present elements. There are various ways of measuring situation awareness, e.g., self-ratings or expert-observer ratings (Endsley & Garland, 2000; Matthews, Eid, Johnsen, & Boe, 2011).

The following discussion provides more information about the four cognitive readiness attributes that are depicted in Table 1.1 and Fig. 1.1:

1. Adaptive expertise
Hatano and Oura (2003) made a distinction between two types of experts—"routine experts" and "adaptive experts." Routine experts have had years of problem-solving experiences in a given domain and can solve familiar types of problems quickly and accurately, but often fail to go beyond procedural efficiency.

However, adaptive experts can go beyond the routine competencies and can be characterized by their flexible, innovative, and creative competencies within a novel domain, rather than in terms of speed, accuracy, and automaticity of solving familiar problems. Ericsson (this volume) notes that deliberate practice over a 10-year time interval is needed to develop adaptive expertise and for developing adaptive expertise, one would provide learning environments that offer variability in encountered tasks and situations (Ericsson, this volume).

2. Creative thinking

Creativity refers to the skills and dispositions needed for generating ideas and products that are (1) relatively novel, (2) high in quality, and (3) appropriate for the task at hand (Sternberg & Lubart, 1995). Creativity results from developmental interactions among three broad contextual components—personal-psychological attributions, cognitive abilities, and environmental-social factors (see Hong, this volume). Following Torrance's (1999) definition of creativity, most investigators view creativity as having four components, that is, fluency (ability to produce a large number of ideas), novel ideas, flexibility (ability to produce or use a variety of approaches), and elaboration (ability to fill in details) (e.g., O'Neil, Abedi, & Spielberger, 1994). This view of creativity is domain general or independent.

Creativity and innovation are often mentioned interchangeably. However, distinctions between the two concepts are meaningful. Creativity is typically used to refer to the act of producing new ideas, approaches, or actions, while innovation is the process of both generating and applying such creative ideas in some specific context.

Further there is an evolving conceptualization that creativity could be viewed as domain independent vs. domain specific. For example, the Ariel Real Life Problem Solving (Hong & Milgram, 2008) provides people with the opportunity to utilize domain-specific creative thinking ability in a wide variety of specific real-life situations. For example, Hong (this volume) views specific creativity in the interpersonal domain as solving a problem involving peers. A domain-specific creative thinking peer item would be: "At recess time you see that children are hitting another child in your class. The child feels that the other children do not like him. What would you do if you were in his place? What are all the things that are possible?" Students were asked to generate as many possible solutions to the problem as they could.

The relationship between domain-general and domain-specific creative thinking ability based on the ideational fluency measures has been examined in a few previous studies (e.g., Hong, this volume). The teaching of creativity has been reviewed by Simonton (2012).

3. Metacognition

Metacognition has been defined as "thinking about thinking," referring essentially to an awareness and regulation of one's own thought processes (Mueller-Hanson et al., 2005). Thus, it refers to knowledge about cognition as well as the control and regulation of cognition (Pintrich & Schunk, 2002; Mayer, this volume). We view metacognition as the process to mentally plan and check on one's progress toward a goal. Metacognition contains two components: planning and self-monitoring. Planning means that one must have a goal (either assigned or self-directed) and a plan to achieve the goal (Hong, O'Neil, & Feldon, 2005). Self-monitoring means one needs a self-checking mechanism to monitor goal achievement (Hong et al., 2005). We have used a questionnaire to measure both trait (Marsh, Hau, Artelt, Baumert, & Peschar, 2006) and state metacognition (O'Neil & Abedi, 1996). However, qualitative approaches are more often used in the literature, e.g., discourse analysis (Eldar, Eylon, & Romen, 2012). Azevedo, Johnson, Chauncey, and Burkett (2010) stated that better metacognitive skills lead to better situation awareness and eventually to better decisions.

A metacognition teaching strategy was evaluated via discourse analysis by Eldar et al. (2012) and Miller and Geraci (2011). In addition a meta-analysis by Sitzmann and Ely (2011) reported that goal level, persistence, effort, and self-efficacy were the stronger self-regulation constructs related to learning. Our view is that in general, metacognitive skills are hard to teach. We hypothesize that a cognitive overload issue (Sweller et al., 2011) is present, e.g., trying to teach content as well as a metacognition skill at the same time results in cognitive overload and poor performance.

4. Teamwork
   Teamwork, by its very nature, requires members to interact, often during times of stress, with other people. Many have argued that the most complex and taxing performance demands in typical military operational environments are often associated with the requirement to function as part of a team (Bowers & Cannon-Bowers, this volume). In other words, individuals must be competent in the tasks associated with their individual jobs and also proficient in the competencies required to be an effective team member (Bowers & Cannon-Bowers, this volume). For the purposes of cognitive readiness of individuals, we define teamwork as a trait of the individual that predisposes the individual to act as a team member. These skills include (a) adaptability—recognizing problems and responding appropriately; (b) coordination—organizing team activities to complete a task on time; (c) decision making—using available information to make decisions; (d) interpersonal—interacting cooperatively with other team members; (e) leadership—providing direction for the team; and (f) communication—clear and accurate exchange of information. This definition was based on the work of Bowers and Cannon-Bowers (this volume) and led to the creation of a teamwork questionnaire by O'Neil et al. (2003) and by Marshall et al. (2005). A complementary literature on collaborative learning has been reported in the K-12 literature (Johnson & Johnson, 2009; O'Neil & Chuang, 2008).

1.3 Training and Assessment of Cognitive Readiness

1.3.1 Training Issues

A fundamental assumption about cognitive readiness in this chapter is that the knowledge, skills, and attributes of cognitive readiness are trainable and measurable. They are predictors of effectiveness and, therefore, should be included in routine training and assessment, that is, they are malleable—they can be improved or changed prior to performance in the environment. Since the concept of cognitive readiness derives from environments that are characterized with volatility, the purpose of training is to pursue more effective functioning, despite uncertainty and disorder in a changing world.

O'Neil's Cognitive Readiness Model purposefully adopted features that were trainable and excluded ones that were not. In this model, the cognitive readiness attributes are conceptualized as trainable traits with some traits harder to train with brief interventions as they rely heavily on prior knowledge and extensive practice (e.g., adaptive expertise). For example, situation awareness can be enhanced by practice with feedback in complex, simulated environments (Fletcher, 2004). Metacognition can be improved by exercises designed to increase awareness of self-regulatory processes (Eldar et al., 2012; Mueller-Hanson et al., 2005).

While certain features of cognitive readiness might be more trainable compared to others, an initial step for the training of cognitive readiness would be to discern what those crucial and trainable features are for a specific task. A second step would be to develop the training plan that focuses on the certain decisive features.

1.3.2 Training Strategies

A major strategy for the training of cognitive readiness is twofold: focusing on similarity and dissimilarity. The first one encourages trainees to develop a rule of thumb that has fewer exceptions. The second one focuses on the complexity of the domain and the prevalence of exceptions. It is often not possible to give trainees an exhaustive list of the conditions under which the rules and examples do not apply (Neal et al., 2006). Three potentially effective training strategies for cognitive readiness are discussed as follows.

One training strategy is to expose the trainees to diverse situations like those they will encounter on their jobs. Pulakos et al. (2002) empirically demonstrated a positive link between past experience in adaptive situations and adaptive performance. Experiencing a variety of situations that require changes in action and adjustments to the environment does appear to aid in the adaptation process (Mueller-Hanson et al., 2005). This is also consistent with the idea that adaptive performance is enhanced by gaining experience in similar situations (Pulakos et al., 2000).

Another training strategy is to use worked examples. Such a training strategy has been effective in teaching adaptive problem solving (Kirschner, Sweller, & Clark, 2006; Shen & O'Neil, 2008). Worked examples are useful for two reasons. First, trainees may store these examples in memory and may subsequently recognize these types of exceptions when they encounter them again (Jones & Endsley, 2000). Second, incorporating examples of exceptions into practice may encourage effortful processing, particularly if the exceptions are surprising and trainees have made errors on them (Ivancic & Hesketh, 2000). The authenticity of such worked examples would free up working memory for dealing with unpredictable situations.

A third training strategy is the use of games and simulations. These technologies enable participants to see how their reactions and decisions influence not only a specific process but also the working of the system (Bell & Kozlowski, 2002; De Jong, 2011; O'Neil & Perez, 2008).

More training strategies for specific cognitive readiness components are provided by Fletcher and Wind (this volume). Finally instructional strategies that reflect evidence-based principles will also "work" in the teaching of cognitive strategies. Such strategies are use of feedback (Hattie & Gan, 2011), strategies for self-explanation (Fonseca & Chi, 2011), cooperative learning (Johnson & Johnson, 2009; Slavin, 2011), intelligent tutoring (Graesser, D'Mello, & Cade, 2011), and guided inquiry techniques (Furtak, Seidel, Iverson, & Briggs, 2012; Kirschner et al., 2006).

1.3.3 Assessment Issues

Whereas the training issues for many of the cognitive readiness skills and attributes seem relatively tractable, the assessment issues are challenging. Many questions need to be answered. Which cognitive readiness skills and attributes are not easily trainable and thus difficult to assess? What represents valid expertise in any specific domain that permits one to deal with unpredictable problems? How can we define the fluency of cognitive readiness in a domain? (Fluency means expertise exercised in a flowing manner or with automaticity; Baker & O'Neil, 2003). How can the results of assessment be connected to the predictions of future success? The validity (do they measure the right things?), the reliability (do they measure things right?), and the precision (how closely do they distinguish one unit or individual from another?) for the assessment of cognitive readiness need to be developed (Fletcher, 2004). See Fletcher and Wind (this volume) for some assessment measures of cognitive readiness. An assessment method we have used involves the use of a computer simulation that allows individuals to experience a wide variety of potential situations.

1.3.4 Assessment Strategies

Cognitive readiness has been evaluated through multiple methods. For example, a study about psychological adaptation to extreme environments used questionnaires with questions that incorporated a sliding scale ranging from 0 to 100 and open-ended questions (Wood, Lugg, Hysong, & Harm, 1999). In addition to the questionnaires, verbal debriefings were used to confirm and amplify the quantitative data (Wood et al., 1999). However, we believe that computer simulations are ideal environments for assessing cognitive readiness skills and attributes. Simulations can be used to test, analyze, or train at a lower cost and at a lower risk than in real-world situations. Much of the research on simulation has been conducted in the Department of Defense environments. Thus we will use their definition of simulation.

1.4 Definition of Simulation

The Defense Modeling and Simulation Office of the US Department of Defense defines simulation as a method for implementing a model over time (DoD Directive 5000.70, DoD Modeling and Simulation (M&S) Management, 2012). They view simulation as either live simulation, virtual simulation, constructive simulation, or combinations of the types of simulations (Brooks et al., 2004; Yardley, Thie, Schank, Galegher, & Riposo, 2003). These terms are commonly used in military simulation. First, live simulation is the most realistic of the simulations since real people use real equipment and perform in real venues. Training in live simulations can be dangerous because the participants are at nearly the same risk as in real situations and live simulations in many situations are cost prohibitive. Live simulations are controlled to a point and the scenarios may be scripted, thus reducing some uncertainty and lessening the risk. The rules of engagement for live simulations also prevent the intentional lethal use of force.

Virtual simulation is the second type of simulation. Real people operate simulated equipment in a simulated environment. Early military and civilian virtual simulators were initially used for aircraft instrument training. Virtual simulation has been used in aviation since at least 1929 with the advent of the Link Trainer (Andrews, Nullmeyer, Good, & Fitzgerald, 2008; Pausch, Crea, & Conway, 1992). Today, most US Military aviators receive their basic training in flight simulators instead of in the actual aircraft for advanced training as well as team training. Many military aviation simulators are equipped with three-axis motion and the pilot has full or partial environmental vision outside the cockpit. The simulated view of the world may be controlled to present a day or night environment. Other factors may be added in the form of clouds, fog, and turbulence to increase the realism of the simulation. The complexity of military simulators has reached the point where pilots in separate cockpits may simulate formation flying together or coordinating missions with other units (Andrews et al., 2008). Virtual simulation is attractive to the military because of the ability to train personnel less expensively than through the use of real simulation. It is extremely costly to move ships to specific locations for training purposes.

The last type of simulation is constructive simulation. All elements of this model are simulated including the personnel, equipment, and the environment (Brooks et al., 2004). Tabletop simulations, sandbox war gaming, and computer-only simulations are used to test theories and capabilities without any real-world people or equipment. These simulations are relatively inexpensive and safe and have the added characteristic of efficacy for certain tasks. Constructive simulations may be run as fast and as often as the capabilities of the hardware and software will allow. A drawback to constructive simulation is that it provides no training for individuals using the simulation. We need to discuss the impact of a virtual simulation on cognitive readiness as a case study for our methodology.

1.4.1 Using Virtual Simulation to Train and Assess Cognitive Readiness: A Feasibility Study

We conducted a feasibility study to examine the effects of a virtual surface warfare simulation on various aspects of cognitive readiness for naval officers. The purpose of the study was to determine if a naval officer's cognitive readiness increased after training in a team training simulation which used a very complex surface warfare scenario.

One of the missions of the US Navy is to train officers to operate complex equipment and to lead personnel at sea. Many events in combat at sea are unexpected including unanticipated tactics, new technological capabilities, new applications of existing technologies, and the elements of surprise (Fletcher, 2004). Naval officers must be capable of making split-second life-and-death decisions under conditions of uncertainty and stress. Humanitarian relief efforts and other crises, although experienced many times over the years, still present high degrees of uncertainty. Fletcher (2004) asks, "How, then, do we prepare military personnel for the unexpected, which, by definition, is something we cannot anticipate nor decompose into specific tasks, conditions, and standards for training?" (p. 1). Our answer exists in the use of virtual simulation in the training and assessment of cognitive readiness.

The US Navy provides a continuum of training for officers who will have command authority on surface ships. This continuum ranges from the entry-level officer training for service aboard ships up to senior officers assigned to major commands of naval vessels at sea (e.g., an aircraft carrier). There are many military training programs, classes, and schools that exist to train the naval officers who serve on the US Navy ships or serve on the larger US Coast Guard cutters. The term naval officer in this chapter refers to both officers in the US Navy and the US Coast Guard. The term ships refers to both the US Navy ships and the US Coast Guard cutters.

One level is for department heads aboard ship and is the focus of our feasibility study. The department head course is the second of the four levels of the training continuum and is intended for officers with mid-level responsibilities aboard ships. The department head course is a demanding and comprehensive 5-month course consisting of 40 h of instruction per week. Students in the department head course study surface warfare, antisubmarine warfare, antiair warfare, naval administration, coordinated operations, multi-threat operations, expeditionary warfare, naval simulation systems, safety afloat, compressed battlespace, and littorals (near land).

The simulator that the Navy uses as part of the training and assessment in the department head course is the Multi-Mission Tactical Trainer (U.S. Navy, 2012). This simulation provides students with a virtual simulation of the combat information center of a ship at sea. The Multi-Mission Tactical Trainer is a series of rooms where each room simulates the combat information center of a particular ship. The instructors may set up the rooms to simulate each ship acting as an individual entity or as a group of ships working together. Each virtual combat information center has tactical sensors and command and control stations where officers perform as they would at sea. Each station in the virtual combat information center represents a different combat officer station aboard ship, such as air defense officer, surface warfare officer, and tactical action officer. The instructors monitor the actions of the students performing different scenarios either as a single ship operation or as multiple ship operations against simulated aggressors or enemies. We used the evaluation framework of Kirkpatick and Kirkpatrick (2006) to frame the evaluation design while a taxonomy for learning (Anderson et al., 2001) was used to frame issues of retention and transfer.

There were 54 naval officers in our study. Multiple cognitive and affective cognitive constructs were measured. However, we will only discuss the cognitive readiness results. The cognitive readiness measures were administered before and after the simulator experience. The following constructs were measured: problem solving, domain-specific strategies, an elaboration cognitive strategy, metacognitive or control strategies, teamwork, and creativity. The metacognition scales (control strategies and elaboration) were adopted from the PISA study (Marsh et al., 2006). As is discussed in the Hong chapter (this volume), there are many definitions of creativity. For our study here we adopted the definition by Torrance (1999), as our measurement instrument was based on his theory. According to Torrance (1999), creativity includes flexibility, elaboration, originality, and fluency. The measure he developed, the Torrance Test of Creative Thinking (Torrance, 1999), is the most widely used creativity test. It is a trait measure.

Based on Torrance (1999), Abedi (2002) constructed a 60-item multiple-choice test to measure creativity. The purpose of developing this new instrument was to shorten the amount of time required for the administration and scoring of the Torrance Test of Creative Thinking. The reliability and validity of this test were reported by O'Neil et al. (1994). In our feasibility study, the alpha reliability for creativity subscales ranged from 0.50 to 0.86. The 0.50 reliability was a single scale. For all the other measures, alpha reliabilities were above 0.70 indicating acceptable reliability.

Our measurement approach was to view teamwork as a trait of the individual that predisposes the individual to act as a team member. Teamwork is made up of six teamwork processes. There are six processes: (a) adaptability, (b) coordination, (c) decision making, (d) interpersonal, (e) leadership, and (f) communication. We measured the teamwork processes via a teamwork questionnaire (O'Neil et al., 2003). Retention and transfer were also measured. We expected the biggest effect on transfer.

The domain-specific problem-solving strategies were tested via a modification of Mayer's (1998) retention and transfer questions. The retention question was: Write an explanation of how you solved tactical problems in the Multi-Mission Tactical Trainer scenario you just completed. The transfer question was: Write an explanation of how you would improve the Multi-Mission Tactical Trainer scenario that you just completed in order to train for a complex and unpredictable warfare environment. Both participants and three expert instructors completed these measures. Participants' retention and transfer responses were independently scored by two raters against the experts' responses. For estimating interjudge reliability, we computed kappa which was 0.93 (after truncation) for both retention and transfer, which indicates a high reliability between the two raters (a graduate student and a retired naval captain graduate student).

1.4.2 Results

In general, the means for each metacognition subscale were approximately the same (pretest and posttest) and were not statistically different. Although prior research by O'Neil et al. (2003) indicated acceptable reliability for the teamwork questionnaire in this study, the alpha reliabilities in this feasibility study ranged from 0.48 to 0.79. Many of the scales (e.g., coordination, interpersonal, leadership, and coordination) were below 0.70 and thus will need revision and will not be further discussed.

With respect to creativity, none of the subscales were statistically significant before and after the simulation experience. However, the posttest creativity-elaboration subscale approached statistical significance ($p=0.05$). This creativity elaboration increased significantly after exposure to the Multi-Mission Tactical Trainer. Although the increase was significant, it is not practically significant. In general, there was little effect of the Multi-Mission Tactical Trainer on cognitive readiness.

We also correlated our cognitive readiness measures with our retention and transfer measures. The creativity measures of fluency (–0.29) and flexibility (–0.34) were statistically significant, but they were unexpectedly negative. It may be that the Navy training focuses on "school solutions" and the personnel system favors errorless performance for promotion. Thus the organization may not reward fluency and flexibility.

1.4.3 Discussion of Feasibility Study and Its Implications for Cognitive Readiness Research

In summary, the hypothesis that there would be an increase in cognitive readiness following the Multi-Mission Tactical Trainer simulator experience was in general not supported. Further only a few cognitive readiness attributes were significantly related to retention measures, i.e., creativity measures of flexibility (negatively) and fluency (negatively). In general, the knowledge of participants vs. that of experts was very low, i.e., posttest retention was 24 % and transfer was 9 %. The lack of a relationship of cognitive readiness with transfer was not unexpected given the data, i.e., the average performance for participants was approximately one correct transfer answer compared to experts' 10 answers and further 24 individuals out of 52 received a zero transfer score.

Another explanation for the lack of effects may be our use of trait measures rather than state ones. Traits are viewed as relatively stable across time and consistent in magnitude while states vary in intensity and change over time (Spielberger, 1980). Given the short simulation intervention (1.5 h), we should have measured states rather than traits. The study concentrated on three trait measures because they were feasible, reliable, and reasonably valid. The three measures included self-regulation measures adopted from PISA study (Marsh et al., 2006), a teamwork measure adopted from O'Neil et al. (2003), and a creativity measure adopted from Abedi (2002).

Another hypothesis that might explain why there was little relationship between cognitive readiness and retention and transfer may be because the naval officers were less than optimally motivated. The day in which the Multi-Mission Tactical Trainer retention and transfer posttests were administered was after students had a Multi-Mission Tactical Trainer simulation scenario (one and a half hours), had taken the posttest questionnaire (40 min), and then completed their final exam for the class (2 h). As the participants had chosen the Navy as a career, the final exam was a must pass and thus high stakes. In contrast, participants knew that the cognitive readiness questionnaire and retention and transfer questions were for this research study and thus were low stakes. The research data were not provided to the students' instructors and thus could not affect their grade. In hindsight, small increments in performance on retention and transfer measures, the use of trait vs. state measures of cognitive readiness, and possible negative motivational effects led to an environment that was not a good test of our hypothesis of using a simulation to teach and test cognitive readiness. However, one positive outcome of this study is that we now have reliable and valid measures of retention and transfer that we will use in our future Multi-Mission Tactical Trainer work as well as lessons learned for the measurement of cognitive readiness.

One research-based implication is that more constructs of cognitive readiness should be tested in order to have a better understanding of the effects of cognitive readiness in a surface warfare simulation. For example, adaptability should also be measured to create a better assessment of the effects of cognitive readiness. However, there are very few state measures for most cognitive readiness constructs. Our ongoing research is to create such state measures and to refine the construct of creativity as specific creative thinking. These new measures will be tested in the Multi-Mission Tactical Trainer environment in order to create a better assessment and to investigate the impact of use of simulation to train cognitive readiness.

Acknowledgements The work reported herein was supported at the University of Southern California by the Office of Naval Research under Award No. N00014-08-0126 at UCLA (TMT2) and ONR Award No. N00014-09-C-0813. The work was supported by subcontract from UCLA. The opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not reflect the views of the Office of Naval Research. We would like to thank Rear Admiral Thomas S. Rowden, USN, Captain Jeffery S. Jones, USN, Captain David Monroe, USN (Ret), Captain George Ponselle, USN (Ret), Captain David Davis, USN (Ret), Captain Carol Hottenrott, USN, Dr. William Bewley, and Dr. John Lee of CRESST. In addition we would like to thank Ms. Joanne Michiuye, Mr. Rob Sitter, and Ms. Laquita Stewart for their excellent preparation of this manuscript.

References

Abedi, J. (2002). A latent-variable modeling approach to assessing reliability and validity of a creativity instrument. *Journal of Creativity, 14*(2), 267–276.

Abrami, P. C., Bernard, R. M., Borokhovski, E., Wade, A., Surkes, M. A., Tamim, R., et al. (2008). Instructional interventions affecting critical thinking skills and dispositions: A stage 1 meta-analysis. *Review of Educational Research, 78*, 1102–1134.

Anderson, L. W., Krathwohl, D. R., Airasian, P. W., Cruikshank, K. A., Mayer, R. E., Pintrich, P. R., et al. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. New York: Longman.

Andrews, D. H., Nullmeyer, R. T., Good, J., & Fitzgerald, P. C. (2008). Measurement of learning processes in pilot simulation. In E. Baker, J. Dickieson, W. Wulfeck, & H. F. O'Neil (Eds.), *Assessment of problem solving using simulations* (pp. 273–288). New York: Erlbaum.

Azevedo, R., Johnson, A., Chauncey, A., & Burkett, C. (2010). Self-regulated learning with MetaTutor: Advancing the science of learning with metacognitive tools. In M. S. Khine & I. M. Saleh (Eds.), *New science of learning: Computers, cognition, and collaboration in education*. New York: Springer.

Baker, E. L. (2008). Learning and assessment in an accountability context. In K. E. Ryan & L. A. Shepard (Eds.), *The future of test-based educational accountability* (pp. 277–291). New York: Routledge.

Baker, E. L. (2012). *The chimera of validity*. Paper prepared for the Educational Assessment, Accountability, and Equity: Conversations on Validity Around the World Conference, Teachers College, Columbia University, New York. Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Baker, E. L., & Mayer, R. E. (1999). Computer-based assessment of problem solving. *Computers in Human Behavior, 15*, 269–282.

Baker, E. L., & O'Neil, H. F. (2003, February). *Assessment of distance learning*. Presentation at the ONR/NSF-sponsored workshop "Technology Assessment of Web-Based Learning," Redondo Beach, CA.

Banks, D., Bader, P. K., Fleming, P. J., Zaccaro, S. J., & Barber, H. F. (2001, April). *Leader adaptability: The role of work experiences and individual differences*. Paper presented at the 16th Annual Conference of the Society for Industrial and Organizational Psychology, San Diego, CA.

Bell, B. S., & Kozlowski, S. W. J. (2002). Adaptive guidance: Enhancing self-regulation, knowledge, and performance in technology-based training. *Personnel Psychology, 55*, 267–306.

Bennis, W. G. (2003). *On becoming a leader*. New York: Basic Books.

Bolstad, C. A., Cuevas, H. M., Babbitt, B. A., Semple, C. A., & Vestewig, R. E. (2006). Predicting cognitive readiness of military health teams. *Proceedings of the International Ergonomics Association 16th World Congress on Ergonomics* [CD-ROM], Maastricht, The Netherlands. Amsterdam: Elsevier.

Bowers, C. A., Braun, C. C., & Morgan, B. B. (1997). Team workload: Its meaning and measurement. In M. T. Brannick, E. Salas, & C. Prince (Eds.), *Team performance assessment and measurement: Theory, methods, and applications* (pp. 85–108). Mahwah, NJ: Erlbaum.

Brooks, H., DeKeyser, T., Jaskot, D., Sibert, D., Sledd, R., Stilwell, W., et al. (2004). Using agent-based simulation to reduce collateral damage during military operations. In M. H. Jones, S. D. Patek, & B. E. Tawney (Eds.), *Proceedings of the 2004 Systems and Information Engineering Design Symposium* (pp. 71–77). Charlottesville: Department of Systems and Information Engineering, University of Virginia.

Carnevale, A. P., Jayasundera, T., & Cheah, B. (2012). *The college advantage: Weathering the economic storm*. Washington, DC: Center on Education and the Workforce, Georgetown Public Policy Institute, Georgetown University.

Cohen, M. S., Thompson, B. B., Adelman, L., Bresnick, T. A., Shastri, L., & Riedel, S. L. (2000). *Training critical thinking for the battlefield*: *Volume I*: *Basis in cognitive theory and research*. Cognitive Technologies Technical Report 00-2. Arlington, VA.

De Jong, T. (2011). Instruction based on computer simulations. In R. E. Mayer & P. A. Alexander (Eds.), *Handbook of research on learning and instruction* (pp. 446–466). New York: Routledge.

DoD Directive 5000.70, *DoD Modeling and Simulation* (*M&S*) *Management*. (2012).

Eldar, O., Eylon, B., & Romen, M. (2012). A metacognitive teaching strategy for preservice teachers: Collaborative diagnosis of conceptual understanding in science. In A. Zohar & J. Dori (Eds.), *Metacognition in science education: Trends in current research* (pp. 225–250). Berkeley, CA: Springer.

Endsley, M. R. (1988). Design and evaluation for situation awareness enhancement. In *Proceedings of Human Factors Society 32nd Annual Meeting* (pp. 97–101). Santa Monica, CA: Human Factors Society.
Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors, 37*, 32–64.
Endsley, M. R., & Garland, D. J. (Eds.). (2000). *Situation awareness analysis and measurement*. Mahwah, NJ: Erlbaum.
Etter, D. M., Foster, R. E., & Steele, T. P. (2000). Cognitive readiness and advanced distributed learning. *Crosstalk: The Journal of Defense Software Engineering, 13*, 5–6.
Finucane, M. L., & Lees, N. B. (2005). *Decision-making competence of older adults*: *Models and methods*. Paper presented at the National Research Council Workshop on Decision Making by Older Adults, Washington, DC.
Finucane, M. L., Mertz, C. K., Slovic, P., & Schmidt, E. S. (2005). Task complexity and older adults' decision-making competence. *Psychology and Aging, 20*(1), 71–84.
Fletcher, J. D. (2004). *Cognitive readiness: Preparing for the unexpected* (IDA Document D-3061). Alexandria, VA: Institute for Defense Analyses.
Fletcher, J. D. (2009). Education and training technology in the military. *Science, 323*, 72–75.
Fonseca, B. A., & Chi, M. (2011). Instruction based on self-explanation. In R. E. Mayer & P. A. Alexander (Eds.), *Handbook of research on learning and instruction* (pp. 296–321). New York: Routledge.
Furtak, E. M., Seidel, T., Iverson, H., & Briggs, D. C. (2012). Experimental and quasi-experimental studies of inquiry-based science teaching: A meta-analysis. *Review of Educational Research, 82*(3), 300–329.
Goldberg, S. L. (2012). Psychology's contribution to military training. In J. H. Laurence & M. D. Matthews (Eds.), *The Oxford handbook of military psychology* (pp. 241–261). New York: Oxford University Press.
Graesser, A. C., D'Mello, S., & Cade, W. (2011). Instruction based on tutoring. In R. E. Mayer & P. A. Alexander (Eds.), *Handbook of research on learning and instruction* (pp. 408–426). New York: Routledge.
Griffin, P., McGaw, B., & Care, E. (Eds.). (2012). *Assessment and teaching of 21st century skills*. Dordrecht, Netherlands: Springer.
Hatano, G., & Oura, Y. (2003). Commentary: Reconceptualizing school learning using insight from expertise research. *Educational Researcher, 32*(8), 26–29.
Hattie, J., & Gan, M. (2011). Instruction based on feedback. In R. E. Mayer & P. A. Alexander (Eds.), *Handbook of research on learning and instruction*. New York: Routledge.
Hong, E., & Milgram, R. M. (2008). *Preventing talent loss*. New York: Routledge.
Hong, E., O'Neil, H. F., & Feldon, D. (2005). Gender effects on mathematics achievement: Mediating role of state and trait self-regulation. In A. M. Gallagher & J. C. Kaufman (Eds.), *Gender differences in mathematics* (pp. 264–293). Cambridge: Cambridge University Press.
Ivancic, K., & Hesketh, B. (2000). Learning from errors in a driving simulation: Effects on driving skill and self-confidence. *Ergonomics, 43*, 1966–1984.
Jackson, J., Thoemmes, F., Jonkmann, K., Ludtke, O., & Trautwein, U. (2012). Military training and personality trait development: Does the military make the man, or does the man make the military? *Psychological Science, 23*(3), 270–277.
Johnson, W. L., Friedland, L., Watson, A. M., & Surface, E. A. (2012). The art and science of developing intercultural competence. In P. J. Durlach & A. M. Lesgold (Eds.), *Adaptive technologies for training and education* (pp. 261–286). New York: Cambridge.
Johnson, D. W., & Johnson, R. T. (2009). An educational psychology success story: Social interdependence theory and cooperative learning. *Educational Researcher, 38*, 365–379.
Jones, D. G., & Endsley, M. R. (2000). Overcoming representational errors in complex environments. *Human Factors, 42*, 367–378.
Keren, G., & Bruine de Bruin, W. (2003). On the assessment of decision quality: Considerations regarding utility, conflict, and accountability. In D. Hardma & L. Macchi (Eds.), *Thinking:*

Psychological perspectives on reasoning, judgment, and decision making (pp. 347–363). New York: Wiley.

Kirkpatrick, D. L., & Kirkpatrick, J. D. (2006). *Evaluating training programs* (3rd ed.). San Francisco: Berrett-Koehler.

Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experimental, and inquiry-based teaching. *Educational Psychologist, 41*(2), 75–86.

Klein, G. (2008). Naturalistic decision making. *Human Factors, 50*(3), 456–460.

Klieme, E., Hartig, J., & Rauch, D. (2008). The concept of competence in educational contexts. In J. Hartig, E. Klieme, & D. Leutner (Eds.), *Assessment of competencies in educational contexts* (pp. 3–22). Cambridge, MA: Hogrefe & Huber.

Kozlowski, S. W. J., Gully, S. M., Salas, E., & Cannon-Bowers, J. A. (1996). Team leadership and development: Theory, principles, and guidelines for training leaders and teams. In M. Beyerlein, S. Beyerlein, & D. Johnson (Eds.), *Advances in interdisciplinary studies of work teams: Team leadership* (Vol. 3, pp. 251–289). Greenwich, CT: JAI Press.

Laurence, J. H., & Matthews, M. D. (2012). *The Oxford handbook of military psychology*. New York: Oxford University Press.

Marsh, H. W., Hau, K.-T., Artelt, C., Baumert, J., & Peschar, J. L. (2006). OECD's brief self-report measure of educational psychology's most useful affective constructs: Cross-cultural, psychometric comparisons across 25 countries. *International Journal of Testing, 6*(4), 311–360.

Marshall, L., O'Neil, H. F., Chen, A., Kuehl, M., Hsieh, I.-L., & Abedi, J. (2005). Teamwork skills: Assessment and instruction. In J. M. Spector, C. Ohrazda, A. Van Schaack, & D. A. Wiley (Eds.), *Innovations in instructional technology: Essays in honor of M. David Merrill* (pp. 131–149). Mahwah, NJ: Erlbaum.

Matthews, M. D., Eid, J., Johnsen, B. H., & Boe, O. C. (2011). A comparison of expert ratings and self-assessments of situation awareness during a combat fatigue course. *Military Psychology, 23*(2), 125–136.

Mayer, R. E. (1998). Cognitive, metacognitive, and motivational aspects of problem solving. *Instructional Science, 26*(1), 49–63.

Mayer, R. E. (2002). A taxonomy for computer-based assessment of problem solving. *Computers in Human Behavior, 18*(6), 623–632.

Mayrath, M. C., Clarke-Midura, J., & Robinson, D. H. (2012). Introduction to technology-based assessments for 21st century skills. In M. C. Mayrath, J. Clarke-Midura, D. H. Robinson, & G. Schraw (Eds.), *Technology-based assessments for 21st century skills: Theoretical and practical implications from modern research* (pp. 1–12). Charlotte, NC: Information Age.

McLagan, P. A. (1997). Competencies: The next generation. *Training & Development, 51*(5), 40–47.

Miller, T. M., & Geraci, L. (2011). Training metacognition in the classroom: The influence of incentives and feedback on exam predictions. *Metacognition Learning, 6*, 303–314.

Morrison, J. E., & Fletcher, J. D. (2001). *Cognitive readiness* (IDA Paper P-3735). Alexandria, VA: Institute for Defense Analyses.

Mueller-Hanson, R. A., White, S. S., Dorsey, D. W., & Pulakos, E. D. (2005). *Training adaptable leaders*: *Lessons from research and practice* (Army Project Number 633007A792). Arlington, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Mumford, M. D., Scott, G. M., Gaddis, B., & Strange, J. M. (2002). Leading creative people: Orchestrating expertise and relationships. *The Leadership Quarterly, 13*, 705–750.

Neal, A., Kirkpatrick, T., Dewsnap, G., Godley, S. T., Joung, W., & Hesketh, B. (2006). An examination of learning processes during critical incident training: Implications for the development of adaptable trainees. *Journal of Applied Psychology, 91*(6), 1276–1291.

O'Neil, H. F., Jr. (1999). Perspectives on computer-based performance assessment of problem solving: Editor's introduction. *Computers in Human Behavior, 15*, 255–268.

O'Neil, H. F., Jr., & Abedi, J. (1996). Reliability and validity of a state metacognitive inventory: Potential for alternative assessment. *The Journal of Educational Research, 89*, 234–245.

O'Neil, H. F., Jr., Abedi, J., & Spielberger, C. (1994). The measurement and teaching of creativity. In H. F. O'Neil Jr. & M. Drillings (Eds.), *Motivation: Theory and research* (pp. 245–263). Hillsdale, NJ: Erlbaum.

O'Neil, H. F., & Chuang, S.-H. (2008). Measuring collaborative problem solving in low-stakes tests. In E. L. Baker, J. Dickieson, W. Wulfeck, & H. F. O'Neil (Eds.), *Assessment of problem solving using simulations* (pp. 177–199). Mahwah, NJ: Erlbaum.

O'Neil, H. F., & Perez, R. S. (Eds.). (2008). *Computer games and team and individual learning*. Oxford, UK: Elsevier.

O'Neil, H. F., Jr., Wang, S., Lee, C., Mulkey, J., & Baker, E. L. (2003). Assessment of teamwork skills via a teamwork questionnaire. In H. F. O'Neil Jr. & R. S. Perez (Eds.), *Technology applications in education: A learning view* (pp. 283–303). Mahwah, NJ: Erlbaum.

OECD. (2012). *Equity and quality in education*: *Supporting disadvantaged students and schools*. OECD Publishing. Retrieved from http://dx.doi.org/10.1787/9789264130852-en

Pausch, R., Crea, T., & Conway, M. (1992). A literature survey for virtual environments-military flight simulator visual systems and simulator sickness. *Presence: Teleoperators and Virtual Environments, 1*(3), 344–363.

Pintrich, P. R., & Schunk, D. H. (2002). *Motivation in education: Theory, research, and applications* (2nd ed.). Upper Saddle River, NJ: Pearson Education.

Pulakos, E. D., Arad, S., Donovan, M. A., & Plamondon, K. E. (2000). Adaptability in the workplace: Development of a taxonomy of adaptive performance. *Journal of Applied Psychology, 85*(4), 612–624.

Pulakos, E. D., Schmitt, N., Dorsey, D. W., Arad, S., Hedge, J. W., & Borman, W. C. (2002). Predicting adaptive performance: Further tests of a model of adaptability. *Human Performance, 15*, 299–323.

Schleicher, A. (2006). *The economics of knowledge*: *Why education is key for Europe*'*s success*. Brussels: Lisbon Council. Retrieved from http://www.oecd.org/dataoecd/43/11/36278531.pdf

Schneider, B. (1994). HRM: A service perspective—Toward a customer-focused HRM. *International Journal of Service Industry Management, 5*, 64–76.

Shen, C.-Y. D., & O'Neil, H. F. (2008). Role of worked examples to stimulate learning in a game. In H. F. O'Neil & R. S. Perez (Eds.), *Computer games and team and individual learning* (pp. 185–204). Oxford, UK: Elsevier.

Simonton, D. K. (2012). Teaching creativity: Current findings, trends, and controversies in the psychology of creativity. *Teaching of Psychology, 39*, 203–208.

Sitzmann, T., & Ely, K. (2011). A meta-analysis of self-regulated learning in work-related training and educational attainment: What we know and where we need to go. *Psychological Bulletin, 137*, 421–442.

Slavin, R. (2011). Instruction based on cooperative learning. In R. E. Mayer & P. A. Alexander (Eds.), *Handbook of research on learning and instruction* (pp. 344–360). New York: Routledge.

Spielberger, C. D. (1980). *Test anxiety inventory*. Palo Alto, CA: Consulting Psychologists Press.

Spiering, B. A., Walker, L. A., Hendrickson, N. R., Simpson, K., Harman, E. A., Allison, S. C., et al. (2012). Reliability of military-relevant tests designed to assess soldier readiness for occupational and combat-related duties. *Military Medicine, 177*(6), 663–668.

Sternberg, R. J., & Lubart, T. I. (1995). *Defying the crowd: Cultivating creativity in a culture of conformity*. New York: Free Press.

Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive load theory*. New York: Springer.

Torrance, E. P. (1999). *Torrance test of creative thinking: Norms and technical manual*. Beaconville, IL: Scholastic Testing Services.

TRADOC Pam 525-8-2. (2011, January 20). The US Army Learning Concept for 2015.

Tucker, J., & Gunther, K. (2009). The application of a model of adaptive performance to Army leader behaviors. *Military Psychology, 21*(3), 315–333.

U.S. Navy, Surface Warfare Officers School Command. (2012). *Multi-Mission Tactical Trainer* (*MMTT*), *Tactical Action Officer* (*TAO*) *Intelligent Tutoring System* (*ITS*). Retrieved from https://www.netc.navy.mil/centers/swos/Simulators.aspx?ID=2

West, R. F., Toplak, M. E., & Stanovich, K. E. (2008). Heuristics and biases as measures of critical thinking: Associations with cognitive ability and thinking dispositions. *Journal of Educational Psychology, 100*(4), 930–941.

Wong, L. (2004). *Developing adaptive leaders: The crucible experience of operation IRAQI freedom*. Carlisle, PA: Strategic Studies Institute, U.S. Army War College.

Wood, J., Lugg, D. J., Hysong, S. J., & Harm, D. L. (1999). Psychological changes in hundred-day remote Antarctic field groups. *Environment and Behavior, 31*(3), 299–337.

Worthy, D. A., Gorlick, M. A., Pacheco, J. L., Schnyer, D. M., & Maddox, W. T. (2011). With age comes wisdom: Decision making in younger and older adults. *Psychological Science, 22*(11), 1375–1380.

Yardley, R. J., Thie, H. J., Schank, J. F., Galegher, J., & Riposo, J. L. (2003). *Use of simulation for training in the U.S. Navy surface force*. Santa Monica, CA: Rand Corporation.

Zaccaro, S. J., & Banks, D. (2004). Leader visioning and adaptability: Bridging the gap between research and practice on developing the ability to manage change. *Human Resource Management, 43*(4), 367.

Chapter 2
The Evolving Definition of Cognitive Readiness for Military Operations

J.D. Fletcher and Alexander P. Wind

Readiness, including cognitive readiness, means different things to different people. In education, we speak of readiness to learn—readiness for third-grade reading, readiness for Algebra I, readiness for Physics 101, and so forth. In training, we usually mean readiness to do something—perform a task, job, or mission. Certainly the two meanings are related. They are both intended to be measureable indications of preparation to do something. Learning is itself doing something, and, conversely, the act of doing something may bring about learning. Interdependency between these two perspectives deserves more discussion, but that must remain an issue for elsewhere. The focus in this chapter is on preparation to perform task-based military missions.

2.1 Readiness

Operational readiness in the Department of Defense (DoD) keys on items that can be measured. Historically, these items fall into four basic categories that are reported for each organizational unit:

- Materiel—enough "systems" such as aircraft, tanks, trucks, radios, and radars
- Equipment—enough spares, supplies, and consumables
- Personnel—enough people certified by training to have the necessary skills at the necessary skill levels to perform anticipated unit missions

J.D. Fletcher, Ph.D. (✉)
Science and Technology Division, Institute for Defense Analyses,
4850 Mark Center Drive, Alexandria, VA, USA
e-mail: fletcher@ida.org

A.P. Wind
State University of New York, Albany, NY, USA
e-mail: AW598554@albany.edu

H.F. O'Neil et al. (eds.), *Teaching and Measuring Cognitive Readiness*,
DOI 10.1007/978-1-4614-7579-8_2,

- Training—completion by the unit of a required number of annual training events such as field exercises, firing range exercises, command and control exercises, and fleet exercises

These categories have been used in one form or another for many years and by various mandated reporting systems to assess each unit's readiness to perform its missions. A unit that satisfies criterion levels for materiel, equipment, personnel, and training is assumed to be "ready." The four categories have been found to be useful and necessary for assessing unit readiness. The issue raised by cognitive readiness is whether more could and possibly should be done to assess readiness.

2.2 Cognitive Readiness

Nothing is so certain in military operations as unpredictability. Unanticipated tactics, new technological capabilities, novel applications of existing technologies, and surprise are all notoriously characteristic of combat engagements. Noncombat military operations such as peacemaking, peacekeeping, humanitarian relief, and crisis management are also known for their potential to bring on unanticipated challenges.

Military units and personnel can be prepared to assume anticipated roles and responsibilities, and much can be done to train them for the missions they are expected to perform. These missions can be decomposed into specific tasks. The tasks may be identified as essential for individuals and units to perform alone or with other individuals and units including those involved in joint (multiservice and/or multinational) missions. The tasks are described in detail for readiness assessment along with the conditions under which they are to be performed and the standards for performance that they must satisfy. If all goes well, the resulting lists of tasks lead to education and training objectives.

The reductionist nature of this approach has been a matter of concern because of the eventual need to deal with the whole of unit performance once these tasks are reaggregated into mission functions (e.g., Hiller, 1987). Reassembling tasks into capable mission performance is addressed by the recent development of mission-essential competencies, which tie the successful performance of mission-essential tasks to the underlying cognitive capabilities needed to perform them successfully (e.g., Alliger, Colegrove, & Bennett, 2003; Chapman, Colegrove, & Greschke, in press). But even with task and competency requirements fully met, the potential for chaos, the unexpected, awaits.

How, then, do we prepare people, teams, and organizational units for the unexpected, which by definition is something we cannot anticipate? Following the lead of Etter, Foster, and Steele (2000), we began to treat this matter as an issue of cognitive readiness, for which Morrison and Fletcher (2002) tentatively suggested the following definition:

> *Cognitive readiness* is the mental preparation (including skills, knowledge, abilities, motivations, and personal dispositions) an individual needs to establish and sustain competent performance in the complex and unpredictable environment of modern military operations. (p. I-3)

Military training, especially training provided to certify and ensure personnel readiness, provides instruction in necessary knowledge and skills, which are embodied in instructional content, such as facts, concepts, and straightforward procedures, and instructional objectives, such as memorizing, understanding, and applying this content.

But "training readiness," which is a central issue for local unit commanders, keys on more context-free, transferable instructional abilities that match content such as adaptive procedures with objectives such as analysis, evaluation, and creative synthesis of new approaches. These abilities allow individuals and units to adapt the knowledge and skills they possess to rapidly evolving operational environments. Training to acquire these abilities is based on realistic experiences that develop what Sternberg (e.g., 2006b; Sternberg & Hedlund, 2002) describes as tacit knowledge—knowledge built up from experience that enables individuals to solve real-world, practical problems. This knowledge is also suggested by Klein's discussion (e.g., 1989) of recognition-primed decision-making focused on the rapid, intuitive decisions so frequently required by military operations.

Both tacit knowledge and recognition-primed decision-making are relevant to current notions of cognitive readiness. This is where the authenticated, situated experiences, so strongly promoted by constructivists (e.g., Tobias & Duffy, 2009) and delivered by simulations and games (e.g., Tobias & Fletcher, 2011), may find a critical role. The primary ingredient added by cognitive readiness is the inclusion of unexpected problems and opportunities in training simulations used to develop individual and unit readiness.

The unexpected is a frequent and widespread characteristic of the free-play, force-on-force engagements found in training exercises—and in military operations themselves. Free-play exercises enable leaders at every level to develop abstract, relatively context-free competencies that can be enhanced through training and measured to assess readiness (Chatham & Braddock, 2001; Gorman, 1990). They provide opportunities for the experiences and pattern recognition development upon which Sternberg's tacit knowledge and Klein's recognition-primed Decision-Making can be established at the general, context-independent level needed to deal with the unexpected.

Cognitive readiness then applies to all military operations, but it is particularly relevant to situations that arise suddenly and require immediate attention—situations that are increasingly characteristic of today's "irregular warfare" operating environment. These characteristics, expanded from those listed by Hurley, Resnick, and Wahlman (2007), include:

- Central role of "human terrain" (local culture, language, customs, norms, and mores)
- Close cooperation of civilian and military organizations in performing their operations
- Emphasis on small-unit operations as they engage allied, opposing, and nonaligned individuals and groups
- Consolidation and coordination of peacemaking, peacekeeping, and social reconstruction operations
- Activity to ensure that host nation's military and civilian organizations are sufficiently secure, stable, and legitimate to assume the responsibilities assigned to them
- Abandonment by adversaries of established norms in treatment of prisoners and nonaligned civilians and the use of chemical, biological, radiological, and nuclear weaponry

Cognitive readiness is as relevant to irregular as to regular warfare. In both cases it raises a number of issues. Can it be better defined? Can it be measured? Can it be trained? Should it be routinely included in routine readiness assessments of individuals, teams, and/or units—military and otherwise?

Post-engagement reviews and analyses such as those provided by Christianson and Zirkle (1992), Orlansky and Thorpe (1992), and Knarr and Richbourg (2008) point to a general description of cognitive readiness in military operations as the ability to:

(a) Remove ambiguity and recognize patterns in uncertain, confusing, and chaotic situations.
(b) Identify and prioritize problems and opportunities presented by these situations.
(c) Devise effective responses to the problems or opportunities presented.
(d) Implement these responses.

Given these characteristics, what components or attributes do they, then, suggest for cognitive readiness? To get the conversation started, Morrison and Fletcher (2002) suggested the attributes shown in column two of Table 2.1. Column three lists subsequent components proposed by O'Neil and discussed by O'Neil, Lang, Perez, Escalante, and Fox (this volume). Column four lists the components proposed and reviewed for this chapter.

O'Neil made a significant advance over the original suggestions by Morrison and Fletcher by further reviewing and analyzing cognitive readiness into specific knowledge, skills, and attributes. He eliminated Transfer, Memory, Automaticity, and Leadership from those proposed by Morrison and Fletcher. He also eliminated Emotion, which he deemed an affective, noncognitive component. He peeled off Adaptive Expertise as a competency from Adaptability as a skill and added Teamwork and Communication. Finally his model includes Knowledge, in the form of Prerequisite and Context Knowledge.

Table 2.1 Components of cognitive readiness proposed by Morrison and Fletcher (column two), O'Neil (column three), and this chapter (column four)

Attribute	Morrison and Fletcher	O'Neil	Fletcher and Wind
Situation Awareness	X	X	X
Problem Solving	X	X	X
Metacognition	X	X	X
Decision-Making	X	X	X
Adaptability	X	X	X
Creativity	X	X	X
Transfer	X		
Pattern Recognition	X		X
Automaticity	X		
Leadership	X		
Emotion	X		
Teamwork		X	X
Communication		X	X
Adaptive Expertise		X	
Interpersonal Skills			X
Resilience			X
Critical Thinking			X

This chapter argues that the components of cognitive readiness should be relatively content- and context-free. If they were not, we would be dealing with anticipated matters and no longer in the land of the unanticipated and unexpected. For this reason, "Knowledge" components are not included in the framework presented here.

This suggestion may occasion some debate by echoing the notion of developing cognitive functions rather than specifically targeted knowledge and skills—of learning Latin to develop cognitive functions that will help us to learn German, solve math problems, or perform other cognitive activities. The contrary idea, which reaches back to E. L. Thorndike's transfer through identical elements (e.g., 1913; Thorndike & Woodworth, 1901) is that if the objective is to learn German or solve math problems, then we should study German or math, not begin with Latin—valuable though that may be in keeping classics scholars employed.

However, in dealing with the unexpected we simply do not know in advance what the objectives of instruction should be or what identical elements they may present. We need to find ways to develop widely usable and context-independent abilities or to select individuals who already have them. This suggestion brings us to the candidate components listed in column four of Table 2.1 and proposed for consideration in this chapter.

The components originally suggested by Morrison and Fletcher and included by O'Neil and this chapter are Situation Awareness, Problem Solving, Metacognition, Decision-Making, Adaptability, and Creative Thinking. We accepted O'Neil's elimination of Transfer, Automaticity, and Leadership. We also accepted his elimination of Emotion from the list, although we view and include emotional control as an aspect of Resilience. We agreed with O'Neil's inclusion of Communication and Teamwork, while adding Interpersonal Skills, Resilience, and Critical Thinking. All these components may eventually need to be bundled or unbundled as thought and research continue. We did not separate out Adaptive Expertise, but left it bundled under Adaptability, even though our focus is quite close to O'Neil's notion of Adaptive Expertise.

We retained Memory and Pattern Recognition as components, but, as our thinking progressed, we have now focused almost exclusively on the pattern recognition functions of memory and list Pattern Recognition by itself in Table 2.1. In accord with Wickens and Flack (1988), Mayer (2005), Bolstad, Endsley, and Cuevas (this volume), and others, we view Pattern Recognition as the basis for integrating the sensory information (visual, aural, etc.) introduced by preceptors (eyes, ears, etc.) in working memory with the contents of and patterns included in long-term memory.

Pattern Recognition applies an abduction process based on intention and long-term store to identify, organize, and separate out what matters in sensory input from what does not. We suggest that Pattern Recognition is a rapid cognitive activity that enables and leads to relatively more deliberate and conscious processes of developing Situation Awareness to support Decision-Making. Pattern Recognition then continues to serve the processes required for Situation Awareness, and the two remain interdependent, but, in this framework, different.

We included Problem Solving from the original Morrison and Fletcher components, as does O'Neil, but eliminated Transfer as too context-dependent to include. We recognize that far transfer, where more abstract, analogous reasoning finds application, may be closer to Problem Solving than near transfer, which depends more on specific, concrete elements common to both settings (e.g., Barnet & Ceci, 2002). We also surmise that Transfer resembles Problem Solving more as we ascend from what Salomon and Perkins (1989) described as low-road transfer (which occurs almost automatically, with little, if any, conscious thought) to high-road transfer (requiring conscious thought and meta-cognitive skill).

Finally, Mayer and Wittrock (1996) address the relationship between Problem Solving and Transfer with their contrast of knowledge transfer and problem-solving transfer. They describe knowledge transfer as focused on learning, when prior learning improves speed or accuracy in learning something new. They describe problem-solving transfer as occurring when solving one problem improves speed or accuracy in the act of solving another one. In cognitive readiness for military operations, the issue is particularly tied to speed and accuracy in Problem Solving. The intersection between Transfer, which we eliminated, and Pattern Recognition and Problem Solving, which we include, seems real. Some forms of Transfer may serve to enable both Pattern Recognition and Problem Solving, but these two appear to be the core issues in cognitive readiness rather than Transfer, especially when it comes to dealing with the unexpected.

If attributes of cognitive readiness are to be viewed as indicators of likely operational effectiveness and included in routine assessments of military readiness, then the next step is to examine evidence that these attributes are measureable and trainable. The position of this chapter—and its current contention in furthering the conversation concerning cognitive readiness—is that these components should be:

1. Relatively content- and context-free. Each component must be applicable to a wide variety of situations that may be neither anticipated nor expected.
2. Measurable. A component of cognitive readiness may be valuable or even essential, but if it cannot be defined or detected through measurement, then it must remain a matter of chance—not a quality that is amenable to systematic assessment, selection, and development.
3. Trainable. An attribute of cognitive readiness should be trainable and not just a matter of selecting personnel for military duties. To some degree it may be born and not made, but, as a practical matter, it should be amenable to enhancement and shaping through instruction (training and/or education) to meet the needs of military operations—or any other human activity.

Of these, the third (trainable) may be the most controversial. Cognitive readiness could include components that are not trainable, but if there is nothing to be done to develop and improve them, then the military can only hope its recruiting and selection processes will provide the cognitive readiness it needs. Enhancing the availability of cognitive competencies through training and/or education seems as essential in this area as it is in others.

2.2.1 Discussion of Candidate Components

The following comments address each of the components proposed in the fourth column of Table 2.1 as candidates for inclusion in the concept of cognitive readiness. The components are addressed in the same semi-chronological order that they are presented in the table. Each is briefly described along with equally brief discussions of measures for them and evidence that they can be improved through instruction.

Situation Awareness. This component of cognitive readiness is the deliberate process based on Pattern Recognition needed to identify in any current situation what elements are relevant for achieving mission goals and to project from that how they will evolve. Endsley (1998, 2006) has provided a three-level definition of Situation Awareness that can be described as (1) the perception of elements in the environment within a volume of time and space, (2) the comprehension of their meaning, and (3) the projection of their status in the near future. Situation Awareness thereby provides the perceptual analysis that precedes decision and action.

Measurement. Evidence presented by a variety of researchers suggests that Situation Awareness can be successfully measured by the "freeze method," which involves stopping ongoing activity and examining participants' perceptions and understanding of it. Salmon et al. (2009) compared the freeze method with a posttrial subjective report and found that the scores on the former were the only measures that significantly correlated with performance on a complex task. Endsley (1995) found that it did not intrude on behavior in performing tasks and that its validity was not affected by the forgetting that can limit post-task reflection. The Situation Awareness Global Assessment Technique (SAGAT) (Endsley, 2000) is a systematic assessment using the freeze method that has been found to be both reliable and valid.

Instruction. A number of sources report that Situation Awareness can be improved through training. For instance, Soliman and Mathna (2010) found that training in modeling as a meta-cognitive strategy significantly improved scores on the SAGAT test and reduced infringements in an automobile driving task. Saus et al. (2006) found that training in reflection using the freeze method improved measures of Situation Awareness in a police shoot-not-shoot simulator. Endsley and Robertson (2000) discussed training to enhance the Situation Awareness of teams, noting that more work needs to be done in this area.

Problem Solving. Mayer and others (e.g., Mayer, 2008; Mayer & Wittrock, 1996) have defined Problem Solving as an effort to achieve a goal by transforming a given situation into an objective situation when it is not immediately obvious how to do that. Early on, Miller, Galanter, and Pribram (1960) demonstrated that Problem Solving can be cast as an analysis of task goals and subgoals. Later, Newell and Simon (1972) showed how means-ends analyses can be used to transform this analysis into a plan of action. Baker and Mayer (1999) characterized Problem Solving as cognitive, process-based, goal directed, and dependent on the capabilities of the problem solver.

Problem Solving has been the object of research from the beginning of experimental psychology. It is a multifaceted, complex area of investigation that has produced an equally multifaceted, complex literature. It is fortunate that for cognitive readiness, we need only to decide if a generalized, relatively context-free capability for Problem Solving can be identified, measured, and taught.

Measurement. Baker and Mayer (1999) distinguish between diagnostic measures of the cognitive processes used to solve problems and outcome measures of success in assessments of Problem Solving. Self-report measures such as the Problem Solving Inventory, which has been in wide use for over 20 years, are generally, if not entirely, focused on processes (D'Zurilla & Nezu, 1990). It and others of the same sort are primarily used for counseling—assessing, ameliorating, and preventing psychological disorders. As D'Zurilla and Maydeu-Olivares (1995) and others emphasize, self-report measures of Problem Solving do not assess this skill directly and may lack empirical support for their construct validity.

Computer-based assessments appear to offer considerable promise for Problem Solving in that they can present problems more realistically, manipulate difficulty in real time, and collect both process and outcome data (Baker & Mayer, 1999; O'Neil, 1999, 2002). Chung, de Vries, Cheak, Stevens, and Bewley (2002) linked participants' verbal protocols with their click-stream data to assess problem-solving ability with IMMEX (Interactive Multimedia Exercises), for which they found substantial evidence of validity. If their results can be found in other contexts, measurement of general, context-free Problem Solving may find reliability and validity though interactive computer-based assessments.

Instruction. Techniques such as focusing on problem definition and analysis, acquiring meta-cognitive skills, and managing mental set have been shown to help individuals improve their skills in solving practical problems (Baker, Niemi, & Chung, 2008), but the large research base on teaching Problem Solving mostly concerns specific contexts. Findings from efforts to teach general-purpose, relatively context-free Problem Solving are mixed and rare (Halpern, 2002; Mayer & Wittrock, 1996).

Reed (1987) found that students typically failed to recognize structural features used to solve problems in one subject area that could be directly applied to solve problems in others. Chen (2010) found students instructed in Problem Solving through a database search exercise showed no transfer on other problem solving metrics. On the other hand, Bloom and Broder (1950) found that college students who were taught problem-solving processes such as analyzing problems into parts and comparing their problem-solving processes to those of experts, scored higher on college examinations than students who did not receive the instruction. Kalyuga and Hanham (2011) used cognitive load theory to study flexible problem-solving skills for applying knowledge structures to new situations and found that these skills can be improved by providing instruction in the generalized forms of knowledge structures. In their review of knowledge transfer and problem-solving transfer, Mayer and Wittrock (1996) concluded that instruction that directly focuses on teaching for problem-solving transfer can improve the general problem-solving skills of students.

Metacognition. Metacognition refers to the executive functions of cognition, particularly those pertaining to knowledge and regulation of one's cognitive processes. Hacker (2001) emphasized the self-regulatory characteristics of Metacognition by noting that it enables individuals to be "agents of their own thinking" (p. 50).

Metacognition contrasts in an interesting way with Klein's (2003) notions of intuitive Decision-Making and Sternberg's practical intelligence (Sternberg & Hedlund, 2002) in that it is often viewed as the capacity to bring an automated (unconscious) skill under conscious control, making individuals aware of their own cognitive processes during task performance. The balance of automatic, intuitive Decision-Making with systematic, deliberate responding remains an important but poorly understood aspect of timely Decision-Making and action in cognitive readiness.

Measurement. Metacognition has been measured by self-report questionnaires such as multiple-choice Self-Regulated Learning scales (Hong, O'Neil, & Feldon, 2005), and the Metacognitive Awareness Questionnaire (Schraw & Denison, 1994). Especially notable is Students' Approaches to Learning (SAL), a self-report measure of self-regulated learning strategies, self-beliefs, motivation, and learning preferences developed with excruciating care and thoroughness for the Organisation for

Economic Co-operation and Development (OECD) (Marsh, Hau, Artelt, Baumert, & Peschar, 2006). The SAL aimed for a balance between brevity and psychometric reliability and validity. It consists of 52 items and takes about 10 min to administer. About 17 of its items cover meta-cognitive strategies of planning, monitoring, and self-regulation. An assessment using confirmatory factor analysis with a sample of 4,000 15-year-old respondents from 25 countries validated its factor structure.

Other measures involve verbal protocols that ask participants to think out loud as they solve problems. These tasks may include embedded incongruities that participants are expected to report. Alternatively, their performance may be monitored by latencies, under the assumption that awareness of incongruities may increase reaction times (Royer, Cisero, & Carlo, 1993). The Metacognitive Knowledge Monitoring Assessment (KMA) procedure has been investigated in 26 studies by Tobias and Everson (2009) who found it to demonstrate substantial construct validity. Meta-cognitive measures are reviewed and discussed more extensively in a volume edited by Hacker, Dunlosky, and Graesser (2009). Overall, it appears that Metacognition and meta-cognitive process can be measured with sufficient reliability and validity for a wide spectrum of purposes.

Instruction. Instruction can improve meta-cognitive performance in a variety of subjects as concluded by Clark and Wittrock (2000) and Hacker et al. (2009). Clark and Wittrock emphasize that although meta-cognitive skills can be trained, "they require more than a few days of a study-skill course" (p. 75). Mayer and Wittrock (1996) list a number of studies that found context-specific improvement of meta-cognitive skills through instruction. Hartman and Sternberg (1993) identified four approaches—general awareness; use of meta-cognitive strategies; providing heuristics for planning, monitoring, and evaluation; and rewarding self-reflection—that have been found to successfully improve Metacognition. Mevarech and Kramarski (1997) used a related approach involving meta-cognitive questioning, practice, review, and verification to improve meta-cognitive mathematics skills that might be applied successfully elsewhere. In a review of research on Metacognition, Walker and Schneider (2009) present strong evidence that meta-cognitive skills can be improved through instruction and describe a variety of instructional approaches that have been found through research to do so reliably.

Decision-Making. After World War II, many military decision-making models were based on operations research (Gass & Assad, 2005). These models in turn were based on economic theories of utility maximization, which assumes that users are (1) completely informed about all major courses of action that apply to a given situation, (2) sensitive to differences that distinguish the courses of action, and (3) rational in their choice of courses of action (Slovic, Lichtenstein, & Fischoff, 1988). These assumptions are not unreasonable if planners are given sufficient time to implement the decision-making process in accord with the prescriptive requirements of operations research. However, these assumptions become increasingly untenable in time-critical situations when they must deal with unexpected situations.

Alternatively, Klein (1989) developed a "naturalistic" model of Decision-Making for the real-world, time-sensitive exigencies of military operations. Sometimes described as macro-cognition (Klein et al., 2003), this view focuses on rapid, real-world, and "satisficing" (satisfactory and sufficient, but not necessarily optimal, per Simon, 1956) decisions made in response to experience-developed patterns in complex, high-stakes, exigent situations with ill-defined and often multiple goals. It is closely related to Sternberg's notion of tacit knowledge, which is acquired through a mixture of formal training and wide practical experience (Hedlund et al., 2003; Sternberg, 1997; Sternberg & Hedlund, 2002). As these authors point out, experts frequently use macro-cognition and tacit knowledge to solve pressing, complex problems even though they cannot articulate precisely what it is, how they acquired it, or how they have applied it.

In Klein's recognition-primed Decision-Making, a decision-maker initially assesses the situation to recognize familiar patterns stored in memory. Once the decision-maker selects or generates a candidate alternative based on this Pattern Recognition, he/she mentally simulates its implementation in the present situation. If the outcome is acceptable, or satisficing, it is selected and implemented. If the outcome is not acceptable, the decision-maker discards the alternative and generates another. Klein's model, then, differs from operations research in at least two ways. First, its decisions are made rapidly, or intuitively (Klein, 2003), rather than through carefully deliberated selection of alternatives along with their estimated payoff and probability of success. Second, its decisions are only expected to satisfy current needs satisfactorily; they are not necessarily optimal.

Decision-Making based on the operations research model can be learned and applied academically. Because this chapter emphasizes Decision-Making in military operations, it is focused on matters like Klein's recognition-primed Decision-Making and Sternberg's tacit knowledge. The central issue here is whether intuitive Decision-Making, which seems essential for dealing with the unexpected at the tactical and operational level of military activity, can be measured and improved by instruction.

Measurement. Miller (1993) used the Miller Intuitiveness Instrument and its self-reporting approach to assess the use of intuition in Decision-Making. On the basis of this assessment, she found the instrument to be both reliable and valid in quantifying the self-perception of intuitiveness among practicing nurses (Miller, 1993). Kahneman and Klein (2009) discussed the boundary conditions between true intuitive skill and overconfident, biased impressions. They determined that the distinction keys on the predictability of the situation and the decision-maker's opportunities and ability to learn its regularities. They conclude that self-reported experience does not provide valid measures of intuitive decision-making capability. Results on this point are therefore mixed.

In general, research on measuring the accuracy and use of intuition in Decision-Making appears scarce. Such a measure might well be developed by applying Simon's (1992) view that situations provide cues to solutions that experts recognize and apply—perhaps automatically and unconsciously, perhaps through patterns they have stored in long-term memory. Such recognition requires at least three conditions: (1) the situation provides necessary cues; (2) the decision-maker recognizes them as relevant and important (solving the abduction problem); and (3) the decision-maker's memory contains information, such as patterns, concerning these cues. An instrument might be constructed around these conditions.

Instruction. Given that tacit knowledge can be measured (e.g., Hedlund et al., 2003; Sternberg & Hedlund, 2002) and attributed to experience, it seems likely that it can be improved and its acquisition accelerated through experience-providing instruction (e.g., Fletcher & Morrison, 2012). In doing so, it may well enable the recognition-primed decision-making advocated by Klein (2003). Context-specific evidence can be derived from the US military combat training centers that developed out of the Vietnam-era "Top Gun" experience. These centers provide free-play, force-on-force exercises that dramatically improve success in military operations (Chatham & Braddock, 2001; Fletcher, 2009). Given the success of these centers in preparing individuals, teams, and units for a wide variety of military operations, which almost inevitably include the unexpected, it seems likely that context-independent intuitive Decision-Making is improved through instruction of this sort.

Sternberg's measures of Tripartite Intelligence demonstrate that measurable, context-independent individual differences exist in the ability to acquire and apply the tacit knowledge (e.g., Sternberg, 2006b) needed for intuitive Decision-Making. It therefore seems likely that context-independent cognitive readiness for intuitive Decision-Making can be developed and improved through instruction. However, demonstrating that context-independent intuitive Decision-Making can be developed or improved through instruction remains hampered by the absence of a direct measure for it.

Adaptability. Sometimes referred to as cognitive agility, the ability to deal adaptively with unanticipated situations is essential in many contexts and has been discussed by a number of commentators (e.g., Burns & Freeman, 2008, 2010; Morrison & Fletcher, 2002; Fletcher, 2004; Zaccaro, Weis, Chen, & Matthews, this volume). In today's environment for military operations, new adversaries continually arise using new tactics. In academia, students and instructors have to adapt knowledge that is developing at an accelerating rate. In the workforce, military and, otherwise, technological innovations continue to create new occupations, new techniques, and new organizational structures. Adaptive adjustment to these challenges, especially those that are unexpected, is an imperative for individuals and organizations in all sectors.

As Zaccaro, Weis, Chen, and Matthews (this volume) point out, a number of commentators distinguish between routine and adaptive expertise. This is a helpful distinction; it is the latter expertise that is the focus of this chapter. Notably, Zaccaro et al. emphasize the importance of noncognitive contributors to adaptive expertise. These contributors are not considered in this chapter, which focuses on cognitive issues, but they are not insignificant and deserve attention as readiness issues expand into noncognitive areas.

Measurement. Le Pine, Colquitt, and Erez (2000) developed a computerized task with 75 problems in which the rules determining correctness change unexpectedly. They found strong and reliable differences among subjects in their ability to adapt to the changes. Potosky and Ramakrishna (2002) developed a self-report measure of adaptive self-efficacy that included the individuals' belief that they could learn new skills quickly. Finally, reliability and construct validity was found for Sternberg's (1997, 2006) tripartite model of intelligence, which includes Adaptability. Given these results, it seems likely that Adaptability or at least Adaptive Expertise can be measured, if it is not being done already.

Instruction. Haynie's (2005) dissertation study found that Adaptability and functioning in dynamic environments are enabled through meta-cognitive awareness, which in turn can be enhanced through instruction. Research reported by Sloutsky and Fisher (2008) with 4- and 5-year-old children found that Adaptability relied on fairly ordinary mechanisms grounded in associative and attention learning. Enhancing Adaptability (or cognitive agility) through instruction appears likely, but it remains an open question. Fortunately, a number of strategies have been suggested to develop and/or enhance the Adaptability of individuals and teams. Five of these are listed by Zaccaro et al. (this volume): self-regulated training, active learning, error management training, experiential variety, and developmental work experiences. One might also extract candidate instructional strategies from the taxonomy proposed by Pulakos, Arad, Donovan, and Plamondon (2000) for teaching. There may well be other sources. Research remains to verify these and other promising strategies for instruction in this area, but it seems likely that such research will provide the verification it seeks.

Creativity. However adaptive and decisive a leader may be, the ability to produce and implement innovative, nonobvious responses to both expected and unexpected situations remains critical. Creativity, i.e., Creative Thinking, is especially important for understanding and responding to the ill-structured problems presented by military operations where surprise is at a premium.

Measurement. Sternberg's Triarchic Abilities Test (2006b) includes a well-established, reliable, and valid measure of Creativity. It requires open-ended verbal responses in the form of captions for cartoons, written stories, and oral stories, and, as Sternberg emphasizes, it is only weakly correlated with, and therefore separate from, measures of verbal intelligence. Cramond, Matthews-Morgan, Torrance, and Zuo (1999), among others, recommend the Torrance Tests of Creative Thinking, which assess fluency, flexibility, originality, and elaboration, as the best standardized and most widely used measures of Creativity, perhaps because of the extensive body of research on their reliability and validity that has been conducted over time and across different cultures. Other standardized tests of Creativity along with reviews of their psychometric qualities are also available. Current examples would include the Abedi-Schumacher Creativity Test (Auzmendi, Villa, & Abedi, 1996), The Test for Creative Thinking—Drawing Production (Urban, 2004), and the Epstein Creativity Competencies Inventory for Individuals (Epstein, Schmidt, &

Warfel, 2008). It seems reasonable to conclude that Creativity can be measured with some appreciable validity.

However, Hong (this volume) discusses the inconsistencies in current measures of Creativity that arise from corresponding inconsistencies in its definition. Hong addresses the complex interactions of cognition, personal traits, and environment in distinguishing between expert talent and creative talent and their development. Hong also discusses distinctions between context-dependent and context-independent Creativity—the latter being claimed by this chapter for cognitive readiness. On the basis of Hong's review, it seems reasonable to conclude that measuring Creativity as a context-independent capacity is attainable.

Instruction. Sternberg (2006a) discusses an experimental study in which students taught in Creativity-inducing conditions outperformed students in noncreativity conditions. Niu and Liu (2009) found that the Creativity of 180 Chinese high school students was improved through instruction to develop their strategies for Creativity. A number of other studies have also found that measures of Creativity can be increased through instruction. It seems reasonable to conclude that Creativity can be enhanced to an appreciable extent through instruction.

Pattern Recognition. The chaotic nature of military operations and today's irregular warfare ensures that the conditions under which individuals learn tasks will differ from the conditions under which they must perform them. Pattern Recognition is required to abstract from experience, identify the familiar, and distinguish it from the unfamiliar and unexpected. As discussed earlier, Pattern Recognition may provide the basis for Situation Awareness. It may also be the basis for transferring knowledge and skill to new situations. Like Transfer, it may key on the presence of identical elements as argued by Thorndike and Woodworth (1901) who pointed to the presence and necessity of identical elements to ensure successful transfer of what is learned in training to what is needed on the job. The same may hold for the features of Pattern Recognition that enable Transfer.

Two similar views are as follows: (a) the Encoding Specificity Hypothesis (Tulving & Thomson, 1973), which suggests that memory is best when the conditions of memory retrieval are congruent with the conditions of original learning, and (b) Transfer-Appropriate Processing (Morris, Bransford, & Franks, 1977), which suggests that memory performance improves as similarities increase between the processes of encoding and the processes of retrieval. These views further suggest the difficulties of dealing with the unexpected, which in many cases is unfamiliar as well as unanticipated.

These mechanisms appear to key on the process of abduction—the ability to perceive what in a chaotic, complex, and confusing situation is important and what is not. Relatively little research has been devoted to this issue, yet it may be critical in explaining the intuitive Decision-Making that characterizes Klein's "macro-cognitive" Decision-Making (e.g., 2003) and Sternberg's tacit knowledge (e.g., Sternberg & Hedlund, 2002), both discussed earlier.

Measurement. Methods for assessing a basic capability for Pattern Recognition are usually based on spatial representations—presenting a series of patterns, leaving the last element blank, and requiring its completion. These techniques are used as culture- and language-free assessments of intelligence. The focus in cognitive readiness is not on intelligence but directly and perhaps separately on the ability to organize incoming stimuli in accord with information stored in long-term memory. The relationship of this organizing ability to some form of intelligence may be real but incidental to cognitive readiness. Nonetheless measurement of the context-independent ability to recognize patterns from confusing, unreliable, and incomplete perceptual stimuli remains clouded by this relationship. Research is needed to determine how well information from current laboratory tests for Pattern Recognition applies to the real-world problems in the abstract, complex, and confusing arena of military operations.

Much research on Pattern Recognition builds on the early work by Chase and Simon (1973), which showed that chess masters were far superior to novices in recognizing and remembering patterns of play that would actually occur in chess, but about the same as novices when chess pieces had been arranged at random. Fiore, Jentsch, Oser, and Cannon-Bowers (2000) found that experienced Navy pilots were superior to novice pilots in recognizing realistic aircraft instrument patterns, but not in recognizing instrument patterns that would not occur in the actual flight. Similarly, Bilalic, Langner, Erb, and Grodd (2010) compared the visual search performance of chess novices and chess experts in chess-related and chess-unrelated tasks. They found both object recognition and Pattern Recognition to be essential in visual cognition and of likely relevance to explain the mechanisms of everyday perception. Their work suggests ways to assess pattern recognition ability for more context-free settings, but no reliable and valid instrument seems available to measure Pattern Recognition in the real world.

Instruction. As the studies mentioned above suggest, the available data suggest that Pattern Recognition can be learned from experience. To an appreciable extent, the necessary experience can be provided by simulation. However, the relationship between specific aspects of simulation fidelity and their contributions to the development of pattern recognition capabilities is frequently unclear and unavailable for selecting cost-effective levels of fidelity to include in the design of simulators and simulations. Also unclear is the degree to which specific elements of simulations used in training contribute to the development of context-independent Pattern Recognition.

On the basis of currently available studies, it appears that pattern recognition ability can be measured and learned for specific contexts, but the degree to which context-independent Pattern Recognition can be measured or taught is uncertain.

Teamwork. Most military operations involve Teamwork, sometimes with established teams such as crews, but often with pick-up teams requiring rapid establishment of Communication, coordination, and procedures (Salas & Cannon-Bowers, 2000). In both cases, teamwork skills are critical and must be applied across many contexts, sometimes under constraining time pressures.

Bowers and Cannon-Bowers (this volume) identify two areas of competency needed for effective team membership: context-specific task-work and relatively context-independent Teamwork. The focus here on context independence leads to an emphasis on the latter—a concern with the relatively context-independent competencies of individuals that enable them to be effective members of teams. As Table 2.1 shows, we chose to include Teamwork, despite its many dimensions, as a component of cognitive readiness. The competencies of Communication and Interpersonal Skills are also included as separate but Teamwork-related cognitive readiness components. It may make sense to unbundle Teamwork into other components as thought and research on the topic continue.

Measurement. Spies, Carlson, and Geisinger (2010) list about a dozen possibly relevant tests of Teamwork that now await evaluation for their psychometric qualities. O'Neil, Wang, Lee, Mulkey, and Baker (2003) and Marshal et al. (2005) discuss the development of a framework for assessing Teamwork and subsequent development assessment of a promising Teamwork Skills Questionnaire, which is intended to assess an individual's potential to participate successfully as a member of a team. This instrument was developed as a cost-effective alternative to requiring an individual to participate and be observed as a member of a team, which, in addition to being costly and time-consuming, would likely be domain-specific and team-specific. O'Neil et al. and Marshal et al. report a number of studies, including confirmatory factor analyses, that provide substantial evidence for the reliability and content validity of this self-report instrument and for teamwork skill as a trait.

An additional subcomponent of Teamwork might be characterized as social Situation Awareness. Another subcomponent might be what Bray (1982) described as the "assembly effect"—the ability to assemble the right mix of people to comprise a team (not simply getting the best people). This latter issue may more directly concern an individual's capabilities in creating and developing teams than in participating as a team member. Other subcomponents of Teamwork may come up as work on cognitive readiness continues.

Instruction. Reviews such as those by Salas, Rozell, Mullen, and Driskell (1999) found mixed results in preparing teams for the unexpected. But a number of studies, such as that by Marks, Zaccaro, and Mathieu (2000), using carefully focused research on teamwork skill training have found it to be effective in achieving its objectives of enhancing Teamwork and improving team performance. Two examples follow.

Pritchard, Bizo, and Stratford (2006) abstracted five common elements from a thorough review of current teamwork definitions. They were common goal(s), member interdependency, dynamic exchange of information, coordination of task activities, and structuring of team member roles. Pritchard et al. compared the team performance of college students in two cohorts, one of which received teamwork skills training prior to a collaborative learning exercise and one of which did not. Students in the cohort who received the pre-instruction training scored higher in posttest self-evaluations of teamwork skills, self-reported assessments of team cohesion, and performance skills to be learned in the collaborative learning exercise.

Of particular relevance to the teamwork required by military operations are the findings of Ellis, Bell, Ployhart, Hollenbeck, and Ilgen (2005) in research on training used to enhance the effectiveness of 65 four-person action teams in a dynamic command and control simulation. About half of the participants were selected at random to receive generic teamwork training after which all participants were assigned to action teams. Team members were required to monitor a geographic area and make decisions needed to defend it against hostile air and ground intrusions during a 1-h session. In accord with methods and measures recommended by Bowers, Salas, Prince, and Brannick (1992), the task required planning and coordination of independently performed tasks, collaborative problem solving, and communication accompanied by strict control over extraneous variables. Ellis et al. found that the teamwork training increased both declarative knowledge of teamwork and necessary team competencies. They also found that it increased proficiency in planning, task coordination, collaborative Problem Solving, and Communication. Finally, they found that their measures of declarative knowledge of Teamwork were correlated with measures of team performance.

In sum, the context-independent training in teamwork provided in both studies improved the teamwork knowledge of individuals and the performance of teams to which they were assigned. It suggests that carefully designed training in teamwork skills can improve the ability of individuals to perform successfully in teams, but more conclusive evidence awaits further research and investigation.

Communication. Although communication skills are related to and often intertwined and bundled with Interpersonal Skills, we treat them separately in this chapter. We agree with O'Neil on the inclusion of Communication, both written and spoken, as a separate component of cognitive readiness. Articulating messages that are reliably received and well understood appears to be a cognitive skill that is essential for Teamwork and success in the conduct of military operations. For instance, Olmstead (1992) found that communication capabilities accounted for about half of the variance in the performance of command and control teams. We focused on verbal communication in this discussion and did not include visual or other nonverbal communication (e.g., gestures, body language, facial expressions, even clothing), but all forms of communication may later prove to be reasonable candidates for cognitive readiness.

Measurement. Communication seems essential in establishing the shared mental models needed for successful Teamwork. Rubin (1985) found the Communication Competency Assessment Instrument (CCAI), which uses structured observations of communication behavior rather than self-reporting, to be reliable and valid. Many communication measures are used to assess physician-patient communications. For example, the SEGUE (Set the stage, Elicit information, Give information, Understand the patient's perspective, and End the encounter) has been found to produce reliable and valid measures through the use of a behavioral checklist (Makoul, 2001).

Instruction in Communication. Similarly, a number of studies have found that communication skills can be improved through instruction. For instance, Rubin, Welsh, and Buerkel (1995) found that high school students improved their scores on the CCAI by taking a communication class. Perrea, Mohamadou, and Kaur (2010) found self-assessment and peer feedback to be successful in improving communication skills.

Interpersonal Skills. These skills are treated separately from communication skills and teamwork skills although all three are interdependent. They concern the ability to relate to and deal with others, regardless of social or cultural background, especially, but not exclusively, for purposes of communication, coordination, and cooperative effort. They key on an individual's ability to put himself/herself in another's place and another's understanding of an environment or situation. Communication is essential in conveying an individual's understanding to another in order to develop a shared mental model and successful Teamwork. However, it is often the product of interpersonal understanding and skill. Interpersonal Skills involve listening to and understanding others as well as communicating. They involve determining what must be done to accomplish goals that require interactions with others who may or may not be members of a cooperating team.

Measurement. Weitzul (1992) developed research-based guidelines for use in interviews to assess the Interpersonal Skills of an interviewee, but very little else appeared in a search for context-independent measures of Interpersonal Skills. No tests with documented and adequate measures of reliability and validity were found. On the other hand, a search for measures of empathy, roughly putting oneself in another's place, was more successful. Research on empathy appears to be divided between awareness of and reactions to another's internal state. Several instruments have shown substantial reliability and construct validity for measuring empathy and awareness of another's state. A standard and relevant instrument, which has been available for over 30 years and has been assessed a number of times, is the Interpersonal Reactivity Index (Davis, 1980). A recent trend is the use of brain imaging to study empathic reactions (e.g., Montgomery, Seeherman, & Haxby, 2009). There appear to be several reliable and valid measures to use in assessing empathy as it might apply to Teamwork and cognitive readiness.

Instruction. Role-playing exercises and simulations are commonly used to provide instruction in Interpersonal Skills. Schroeder, Dyer, Czerny, Youngling, and Gillotti (1986) developed a series of videodiscs that successfully used simulations and role-playing exercises to develop the Interpersonal Skills of military leaders. Holsbrink-Engels (1997) found that computer-based simulations and role-playing enhanced Interpersonal Skills development. However, development of Interpersonal Skills through instruction has not been validated through the use of standardized measurement. Several studies using a variety of approaches have shown that empathic intensity (Hooker, Verosky, Germine, Knight, & D'Esposito, 2008) and empathic accuracy (Barone et al., 2005) can be increased through instruction.

Resilience. In discussions with military and civilian K-12 educators, Resilience comes up as a characteristic that is critical for success (Burns & Freeman, 2008, 2010). Applied to cognitive readiness, it is “grit,” a refusal to give up in activities ranging from K-12 classroom instruction to sports to military operations. Bonanno (2004), among others, defines resilient individuals as those who maintain healthy, stable, and productive functioning despite being exposed to highly disruptive, traumatic environments or events. The related concept of hardiness is identified as the basis for Resilience (e.g., Bartone, 1999). It is described as consisting of three interrelated attitudes: commitment to experience, control over situations, and challenge to prevail. Bartone (2007) concludes from research on military leadership that it is best predicted by hardiness.

Fredrickson, Tugade, Waugh, and Larkin (2003) were able to reliably identify resilient individuals and found that they were particularly able to use positive emotions to mobilize the psychological, emotional, and cognitive resources needed to successfully cope with significant catastrophes—in military terms, they were able to perform for themselves the duty of commanders to maintain hope. Further, Fredrickson et al. noted the tendency among resilient individuals to seek and focus on positive aspects and meanings in even the most dire of situations. It could be argued that Resiliency is born, not made, but there is evidence that it is measurable and trainable.

Measurement. Tusaie and Dyer (2004) point out that most research on Resilience keys on mental health and an absence of symptoms, rather than direct measures of Resilience. Two measures for Resilience are the Connor-Davidson Resiliency Scale (2003), which is a 25-item measure of Resilience, and the Resilience Scale for Adolescents (READ) (Hjemdal, Friborg, Stiles, Martinussen, & Rosenvinge, 2006), which both Hjemdal et al. and Soest, Mossige, Stefansen, and Hjemdal (2010) have assessed—the latter with a sample of 6,723 teenage subjects. Both of these instruments appear to produce reliable and valid measures of Resilience. Bartone’s Dispositional Resilience Scale (1989) was identified by Funk (1992) as a reliable and valid measure of hardiness and shown by Bartone (2007) to produce a reliable measure of commitment, control, and challenge in a population of military cadets. The capabilities and psychometric properties of these tests indicate that measures of Resilience and hardiness are available.

Instruction. It also appears that Resilience can be improved through instruction. Grant, Curtayne, and Burton (2009) compared 41 executives who received direct coaching conducted by professional executive coaches with a control group that received no coaching demonstrated. Using 360° feedback, the researchers found higher levels of goal attainment, Resilience, and workplace well-being in the coached, experimental group. Liossis, Shochet, Millear, and Biggs (2009) assessed a blended cognitive-behavioral program. Using self-reports, they found increases in such aspects of Resilience as greater self-efficacy, optimism, and work satisfaction immediately after the program and later in a 6-month follow-up. Measures in both these studies were subjective. Improvement of Resilience through instruction seems likely, but more conclusive research and results are needed. Resilience may need to be decomposed into confidence, control, composure, self-regulation, and the like.

Critical Thinking. Critical Thinking is related to Metacognition when applied to one's own thinking, but we have included it as a separate component of cognitive readiness because it may also be applied to the cogitation of others as well as to one's own. When experienced decision-makers deal with complex and unexpected situations they employ Situation Awareness to collect evidence, use abductive processes to select what is relevant, look for patterns consistent with their experience, and apply critical thinking skills to identify and evaluate alternative approaches that may not be optimal, but that satisfice (Simon, 1956, again). In general, critical thinking skills are used by decision-makers to ensure that they have (a) asked the right question, (b) collected, organized, and assessed relevant data, (c) avoided bias and mind sets, (d) identified and evaluated assumptions, and (e) generated and evaluated appropriate hypotheses (e.g., Halpern, 2002; Sternberg, Roediger, & Halpern, 2006).

Measurement. Ennis (1993) identified nine tests that assess more than one aspect of Critical Thinking and four others that assess at least one aspect. More recently, Sobocan and Groarke (2009) provided a useful edited volume on the measurement of Critical Thinking. Our review of critical thinking measurement suggests that the most widely used test is the Watson-Glaser Critical Thinking Appraisal followed next by the Minnesota Test of Critical Thinking. Although Fawkes et al. (2002) identified flaws in 66 of the 80 questions on the Watson-Glaser test, Gadzella et al. (2006) found it to be sufficiently reliable and valid for several different subgroups totaling 586 university students. An extensive assessment by Bernard et al. (2008) concluded that the Watson-Glaser test measured Critical Thinking overall, but that its subscales should not be interpreted individually. Given the number of tests identified by Ennis and the conclusions from Bernard et al., and despite reservations by Fawkes et al., it appears that Critical Thinking can, to some useful extent, be measured.

Instruction. Other studies using one or another critical thinking test found that Critical Thinking can be taught. For example, a meta-analysis by Allen, Berkowitz, Hunt, and Louden (1999) concluded that instruction in forensics, debate, public speaking, argumentation, and the like improved critical thinking ability as measured by critical thinking tests. Dale, Ballotti, Handa, and Zynch (1997) found that a problem-solving course for Purdue freshmen increased scores on the Watson-Glaser test from the 20th to the 35th percentile. Johnson, Flagg, and Dremsa (2007) found simulation to be superior to direct instruction in raising scores on higher level cognition measures including those involving Critical Thinking. On the basis of these and other studies, it seems reasonable to conclude that Critical Thinking can be improved through instruction.

Table 2.2 Summary of findings on proposed cognitive readiness components

Candidate component	Measurement	Instruction
Situation Awareness	Can be reliably measured by the "freeze method"—stopping a task and asking about the situation	Training in reflection and modeling as meta-cognitive strategies found to increase scores on tests of Situation Awareness
Problem Solving	Several self-report tests reliably measure Problem Solving process factors. Content validity remains an issue, but computer-based assessment offers serious promise for validity	Efforts to teach relatively context-free Problem Solving are rare, but instruction that directly focuses on problem solving transfer shows promise
Metacognition	A number of carefully assessed tests and verbal protocols have been found to be reliable and valid	Promoting general awareness, heuristics, and self-reflection have raised scores on various tests of Metacognition
Decision-Making	No standardized, objective measurement with established reliability and validity exists—evidence suggests one could be developed	Experience and simulation promote the background needed, but absence of good measures limits evidence of increased decision making skill through instruction
Adaptability	Reliability and construct validity found for measures that include Adaptability	Evidence exists that Adaptability can be increased along with means to do so, but the issue remains to be more fully determined
Creativity	There appear to be several credible measures of aspects of Creativity	Evidence found that scores on measures of Creativity can be increased through instruction
Pattern Recognition	Laboratory measures exist, but evidence that context-free Pattern Recognition can be measured is still needed	Evidence suggests improvements in Pattern Recognition through instruction but lacks standardized measures
Teamwork	One promising self-report measure of context-independent Teamwork ability was found	Early evidence is mixed, but more recent studies have demonstrated Teamwork improvement through instruction
Communication	Reliability and construct validity found for at least one standard test	Scores on Communication tests have been increased by instruction
Interpersonal Skills	Reliability and construct validity found for empathy, but not for overall Interpersonal Skills	Evidence from simulation and role-playing exercises suggests improvements but lacks standardized measurement. Instruction found to increase empathy
Resilience ("Grit")	Reliability and construct validity found for standardized tests of Resilience	Evidence found of increases through instruction and/or coaching. Resilience may need to be further decomposed
Critical Thinking	Several published tests available. Credible findings exist on their reliability and validity	Studies reported gains in Critical Thinking on standardized tests after instruction

2.3 Summary

A number of candidate components were considered under the criteria that they should be (a) relatively content- and context-free, (b) measurable, and (c) capable of improvement through instruction. These components are briefly reviewed with regard to these criteria in Table 2.2. It is entirely possible that new components that meet these three criteria could be added, others could be combined, and still others, such as Teamwork, Resilience, and Pattern Recognition, may have to be further decomposed. The discussion continues—as it should.

In brief, the six candidate components of cognitive readiness that were found to meet the three criteria we chose for cognitive readiness were Situation Awareness, Metacognition, Creativity, Communication, Resilience, and Critical Thinking. Given the promise of computer-based measurement and instruction focused on context-independent, problem-solving transfer (Mayer & Wittrock, 1996), context-independent Problem Solving seems likely to meet the criteria for measurement and instruction soon. With a large number of measures waiting review for reliability and validity and promising results from performance ratings and surveys, Teamwork also seems likely to meet the criteria soon. The same might be said for Interpersonal Skills based on standardized measures for empathy and research evidence showing improvements from simulation and role-playing exercises. Similar optimism seems justified for Adaptability given Sternberg's (2006b) measures of Adaptability fortified by findings that Adaptability can be enhanced by a focus on meta-cognitive awareness (Haynie, 2005) and by Zaccaro's research for the Army (e.g., 2009). Context-independent Intuitive Decision-Making remains more problematic because it lacks objective measures with demonstrated reliability and validity focused on this capability. It is often seen employed decisively in military operations and in research by Sternberg (e.g., 2006b) and analyses by Klein (e.g., 2003), but empirical work focused on this issue remains needed. Finally, the massive corpus of research and theory on Pattern Recognition may yet yield the standardized measures that we sought for context-independent Pattern Recognition, but none that were appropriate appeared in this survey.

2.3.1 Final Word

The challenge of training individuals and groups of individuals for unexpected situations that cannot by definition be anticipated seems both substantive and of practical significance for behavioral science. Thus far, it appears that our research has at hand relevant and substantial responses to most of the issues raised by cognitive readiness. Opportunities remain along with work to be done, especially if cognitive readiness is to be included in routine assessments of workforce readiness in the military and in industry. It seems to be a promising opportunity and a challenge for behavioral research.

Acknowledgements The work reported herein was partially supported by a grant from the Office of Naval Research, Award Number N000140810126. The findings and opinions expressed here do not necessarily reflect the positions or policies of the Office of Naval Research.

References

Allen, M., Berkowitz, S., Hunt, S., & Louden, A. (1999). A meta-analysis of the impact of forensics and communication education on critical thinking. *Communication Education, 48*(1), 18–30.

Alliger, G. M., Colegrove, C. M., & Bennett, W., Jr. (2003). *Mission essential competencies: New method for defining operational readiness and training requirements*. Paper presented at the Thirteenth International Occupational Analyst Workshop, San Antonio.

Auzmendi, E., Villa, A., & Abedi, J. (1996). Reliability and validity of a newly constructed multiple-choice creativity instrument. *Creativity Research Journal, 9*(1), 89–95.

Baker, E. L., & Mayer, R. E. (1999). Computer-based assessment of problem solving. *Computers in Human Behavior, 15*, 269–282.

Baker, E. L., Niemi, D., & Chung, G. K. W. K. (2008). Simulation and the transfer of problem-solving knowledge and skills. In E. L. Baker, J. Dickieson, W. Wulfeck, & H. F. O'Neil (Eds.), *Assessment of problem solving using simulations* (pp. 1–17). Mahwah, NJ: Erlbaum.

Barnet, S. M., & Ceci, S. J. (2002). When and where do we apply what we learn? A taxonomy of far transfer. *Psychological Bulletin, 128*, 612–637.

Barone, D. F., Hutchings, P. S., Kimmel, H. J., Traub, H. L., Cooper, J. T., & Marshall, C. M. (2005). Increasing empathetic accuracy through practice and feedback in a clinical interviewing course. *Journal of Social and Clinical Psychology, 24*(2), 56–171.

Bartone, P. T. (1999). Hardiness protects against war-related stress in Army reserve forces. *Consulting Psychology Journal: Practice and Research, 51*, 72–82.

Bartone, P. T. (2007). Test-retest reliability of the Dispositional Resilience Scale-15, a brief hardiness scale. *Psychological Reports, 101*, 943–944.

Bernard, R. M., Zhang, D., Abrami, P. C., Sicoly, F., Borokhovski, E., & Surkes, M. A. (2008). Exploring the structure of the Watson-Glaser critical thinking appraisal: One scale or many subscales? *Thinking Skills and Creativity, 3*, 15–22.

Bilalic, M., Langner, R., Erb, M., & Grodd, W. (2010). Mechanisms and neural basis of object and pattern recognition: A study with chess experts. *Journal of Experimental Psychology: General, 139*(4), 728–742.

Bloom, B. S., & Broder, L. J. (1950). *Problem-solving processes of college students*. Chicago, IL: University of Chicago Press.

Bonanno, G. A. (2004). Loss, trauma, and human resilience: Have we underestimated the human capability to thrive after extremely aversive events? *American Psychologist, 59*(1), 20–28.

Bowers, C., Salas, E., Prince, C., & Brannick, M. (1992). Games teams play: A method for investigating team coordination and performance. *Behavior Research Methods, Instrumentation, & Computers, 24*, 503–506.

Bray, D. W. (1982). The assessment center and the study of lives. *American Psychologist, 37*, 180–189.

Burns, W. R., Jr., & Freeman, W. (2008). *Developing an adaptability training strategy and policy for the DoD* (IDA Paper P-4358). Alexandria, VA: Institute for Defense Analyses.

Burns, W. R., Jr., & Freeman, W. (2010). *Developing more adaptable individuals and institutions* (IDA Paper P-4535). Alexandria, VA: Institute for Defense Analyses.

Chapman, R., Colegrove, C. M., & Greschke, D. (in press). Transforming operational training in the combat air forces. *Journal of Military Psychology*.

Chase, W., & Simon, H. A. (1973). *Cognitive Psychology, 4*, 55–81.

Chatham, R. E., & Braddock, J. V. (2001). *Training superiority and training surprise*. Washington, DC: Defense Science Board, Department of Defense.

Chen, C. (2010). Teaching problem solving and database skills that transfer. *Journal of Business Research, 63*(2), 175–181.

Christianson, W. M., & Zirkle, R. A. (1992). *73 Easting battle replication—A Janus combat simulation* (IDA Paper P-2770). Alexandria, VA: Institute for Defense Analyses (DTIC/NTIS No. ADA 257 266).

Chung, G. K. W. K., de Vries, L. F., Cheak, A. M., Stevens, R. H., & Bewley, W. L. (2002). Cognitive process validation of an online problem-solving assessment. *Computers in Human Behavior, 18*, 669–684.

Clark, R., & Wittrock, M. C. (2000). Psychological principles in training. In S. Tobias & J. D. Fletcher (Eds.), *Training & retraining: A handbook for business, industry, government, and the military*. New York: Macmillan Reference.

Connor, K. M., & Davidson, J. T. (2003). Development of a new reliance scale: The Connor-Davidson Resilience Scale (CD-RISC). *Depression and Anxiety, 18*(2), 76–82.

Cramond, B., Matthews-Morgan, J., Torrance, E. P., & Zuo, L. (1999). Why should the Torrance Tests of Creative Thinking be used to assess creativity? *Korean Journal of Thinking & Problem Solving, 9*(2), 77–101.

D'Zurilla, T. J., & Maydeu-Olivares, A. (1995). Conceptual and methodological issues in social problem-solving assessment. *Behavior Therapy, 26*, 409–432.

D'Zurilla, T. J., & Nezu, A. M. (1990). Development and preliminary evaluation of the Social Problem-Solving Inventory. *Psychological Assessment, 2*, 156–163.

Dale, P. M., Ballotti, D., Handa, S., & Zynch, T. (1997). An approach to teaching problem solving in the classroom. *College Student Journal, 31*(1), 76–79.

Davis, M. H. (1980). A multidimensional approach to individual differences in empathy. *JSAS Catalog of Selected Documents in Psychology, 10*, 85–104.

Ellis, A. P. J., Bell, B. S., Ployhart, R. E., Hollenbeck, J. R., & Ilgen, D. R. (2005). An evaluation of generic teamwork skills training with action teams: Effects on cognitive and skill-based outcomes. *Personnel Psychology, 58*, 641–672.

Endsley, M. R. (1995). Measurement of situation awareness in dynamic systems. *Human Factors: The Journal of the Human Factors and Ergonomics Society, 37*(1), 65–84.

Endsley, M. R. (1998). Situation awareness for the individual soldier. In S. E. Graham (Ed.), *Infantry situation awareness workshop*. Fort Benning, GA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Endsley, M. R. (2000). Direct measurement of situation awareness: Validity and use of SAGAT. In M. R. Endsley & D. J. Garland (Eds.), *Situation awareness analysis and measurement* (pp. 131–157). Mahwah, NJ: Erlbaum.

Endsley, M. R. (2006). Situation awareness. In G. Salvendy (Ed.), *Handbook of human factors and ergonomics* (3rd ed., pp. 528–542). Hoboken, NJ: Wiley.

Endsley, M. R., & Robertson, M. M. (2000). Training for situation awareness in individuals and teams. In M. R. Endsley & D. J. Garland (Eds.), *Situation awareness analysis and measurement* (pp. 310–308). Mahwah, NJ: Erlbaum.

Ennis, R. H. (1993). Critical thinking assessment. *Theory into Practice, 32*, 179–186.

Epstein, R., Schmidt, S. M., & Warfel, R. (2008). Measuring and training creative competencies: Validation of a new test. *Creativity Research Journal, 20*(1), 7–12.

Etter, D. M., Foster, R. E., & Steele, T. P. (2000). Cognitive readiness and advanced distributed learning. *Crosstalk: The Journal of Defense Software Engineering, 13*, 5–6.

Fawkes, D., Adajian, T., Flage, D., Hoeltzel, S., Knorpp, B., O'Meara, B., et al. (2002). Examining the exam: A critical look at the Watson-Glaser critical thinking appraisal exam. *Inquiry: Critical Thinking Across the Disciplines, 21*(3), 31–46.

Fiore, S. M., Jentsch, F., Oser, R. L., & Cannon-Bowers, J. A. (2000). Perceptual and conceptual processing in expert/novice cue pattern recognition. *International Journal of Cognitive Technology, 5*(2), 17–26.

Fletcher, J. D. (2004). *Cognitive readiness: Preparing for the unexpected* (IDA Document D-3061). Alexandria, VA: Institute for Defense Analyses (DTIC/NTIS No. ADA 417 618).

Fletcher, J. D. (2009). The military value of expertise and expert performance. In K. A. Ericsson (Ed.), *Development of professional expertise: Toward measurement of expert performance and design of optimal learning environment* (pp. 449–469). Cambridge, UK: Cambridge University Press.

Fletcher, J. D., & Morrison, J. E. (2012). *DARPA Digital Tutor: Assessment Data* (IDA Document D-4686). Alexandria, VA: Institute for Defense Analyses.

Fredrickson, B. L., Tugade, M. M., Waugh, C. E., & Larkin, G. R. (2003). A prospective study of resilience and emotions following the terrorist attacks on the United States on September 11th, 2002. *Journal of Personality and Social Psychology, 84*(2), 365–376.

Gadzella, B. M., Hogan, L., Masten, W., Stacks, J., Stephens, R., & Zascavage, V. (2006). Reliability and validity of the Watson-Glaser Critical Thinking Appraisal Forms for different academic groups. *Journal of Instructional Psychology, 33*(2), 141–143.

Gass, S. I., & Assad, A. S. (2005). *An annotated timeline of operations research: An informal history*. New York: Kluwer.

Gorman, P. F. (1990). *The military value of training* (IDA Paper P-2515). Alexandria, VA: Institute for Defense Analyses (DTIC/NTIS No. ADA 232 460).

Grant, A. M., Curtayne, L., & Burton, G. (2009). Executive coaching enhances goal attainment, resilience, and workplace well-being: A randomized controlled study. *The Journal of Positive Psychology, 4*(5), 396–407.

Hacker, D. J. (2001). *Metacognition: Definitions and empirical foundations* [On-line Report]. Memphis, TN: The University of Memphis. Retrieved May 30, 2010, from http://www.psyc.memphis.edu/trg/meta.htm

Hacker, D. J., Dunlosky, J., & Graesser, A. C. (Eds.). (2009). *Handbook of metacognition in education*. New York: Routledge.

Halpern, D. F. (2002). *Thought & knowledge*. Mahwah, NJ: Erlbaum.

Hartman, H., & Sternberg, R. J. (1993). A broad BACEIS for improving thinking. *Instructional Science, 21*, 401–425.

Haynie, J. M. (2005). Cognitive adaptability: The role of metacognition and feedback in entrepreneurial decision policies (Doctoral dissertation, University of Colorado at Boulder). Retrieved from http://www.dtic.mil/cgi-bin/GetTRDoc?AD=ADA434527

Hedlund, J., Forsythe, G. B., Horvath, J., Williams, W. M., Snook, S., & Sternberg, R. J. (2003). Identifying and assessing tacit knowledge: Understanding the practical intelligence of military leaders. *The Leadership Quarterly, 14*, 117–140.

Hiller, J. H. (1987). Deriving useful lessons from combat simulations. *Defense Management Journal, 23*, 29–33.

Hjemdal, O., Friborg, O., Stiles, T. C., Martinussen, M., & Rosenvinge, J. H. (2006). A new scale for adolescent resilience: Grasping the central protective resources behind healthy development. *Measurement and Evaluation in Counseling and Development, 39*(2), 84–97.

Holsbrink-Engels, G. A. (1997). Computer-based role-playing for interpersonal skills training. *Simulation & Gaming, 28*, 164–180.

Hong, E., O'Neil, H. F., Jr., & Feldon, D. (2005). Gender effects on mathematics achievement: Mediating role of state and trait self-regulation. In A. M. Gallagher & J. C. Kaufman (Eds.), *Gender differences in mathematics* (pp. 264–293). Cambridge, MA: Cambridge University Press.

Hooker, C. I., Verosky, S. C., Germine, L. T., Knight, R. T., & D'Esposito, M. (2008). Mentalizing about emotion and its relationship to empathy. *Social Cognitive and Affective Neuroscience, 3*(3), 204–217.

Hurley, W. J., Resnick, J. B., & Wahlman, A. (2007). *Improving capabilities for irregular warfare, Volume I: Framework and applications* (IDA Paper P-4267). Alexandria, VA: Institute for Defense Analyses.

Johnson, D., Flagg, A., & Dremsa, T. L. (2007). Effects of using a human patient simulator (HPSTM) versus a CD-ROM on cognition and critical thinking. *Medical Education Online,* 13.

Kahneman, D., & Klein, G. (2009). Conditions for intuitive expertise: A failure to disagree. *American Psychologist, 64*(6), 515–526.

Kalyuga, S., & Hanham, J. (2011). Instructing in generalized knowledge structures to develop flexible problem solving skills. *Computers in Human Behavior, 27*(1), 63–68.

Klein, G. A. (1989). Recognition-primed decisions. In W. B. Rouse (Ed.), *Advances in man-machine systems research* (pp. 47–92). Greenwich, CT: JAI Press.

Klein, G. A. (2003). *The power of intuition*. New York: Doubleday.

Klein, G., Ross, K. G., Moon, B. M., Klein, D. E., Hoffman, R. R., & Hollnagel, E. (2003). Macrocognition. *EEE Intelligent Systems, 18*, 81.

Knarr, W., & Richbourg, R. (2008). *Learning from the first victory of the 21st century: Mazar-e Sharif* (IDA Document D-3380). Alexandria, VA: Institute for Defense Analyses.

Le Pine, J. A., Colquitt, J. A., & Erez, A. (2000). Adaptability to changing task contexts: Effects of general cognitive ability, conscientiousness, and openness to experience. *Personnel Psychology, 53*(3), 563–593.

Liossis, P. L., Shochet, I. M., Millear, P. M., & Biggs, H. (2009). The Promoting Adult Resilience (PAR) Program: The effectiveness of the second, shorter pilot of a workplace prevention program. *Behavior Change, 26*(2), 97–112.

Makoul, G. (2001). The SEGUE framework for teaching and assessing communication skills. *Patient Education and Counseling, 45*, 23–34.

Marks, M. A., Zaccaro, S. J., & Mathieu, J. E. (2000). Performance implications of leader briefings and team-interaction training for team adaptation to novel environments. *Journal of Applied Psychology, 85*, 971–986.

Marsh, H. W., Hau, K.-T., Artelt, C., Baumert, J., & Peschar, J. L. (2006). OECD's brief self-report measure of educational psychology's most useful affective constructs: Cross-cultural, psychometric, comparisons across 25 countries. *International Journal of Testing, 6*(4), 311–360.

Marshal, L., O'Neil, H. F., Chen, A., Kuehl, M., Hsieh, I.-L., & Abedi, J. (2005). Innovations in instructional technology: Essays in honor of M. David Merrill. In J. M. Spector, C. Ohrazda, A. Van Schaak, & D. A. Wiley (Eds.), *Innovations in instructional technology: Essays in honor of M. David Merrill* (pp. 131–150). Mahwah, NJ: Erlbaum.

Mayer, R. E. (2005). Cognitive theory of multimedia learning. In R. E. Mayer (Ed.), *The Cambridge handbook of multimedia learning* (pp. 31–48). New York: Cambridge University Press.

Mayer, R. E. (2008). Problem-solving assessment in games and simulation environments. In E. L. Baker, J. Dickieson, W. Wulfeck, & H. F. O'Neil (Eds.), *Assessment of problem solving using simulations* (pp. 139–156). Mahwah, NJ: Erlbaum.

Mayer, R. E., & Wittrock, M. C. (1996). Problem-solving transfer. In D. C. Berliner & R. C. Calfee (Eds.), *Handbook of educational psychology* (pp. 47–62). New York: Macmillan.

Mevarech, Z. R., & Kramarski, B. (1997). IMPROVE: A multidimensional method of teaching mathematics in heterogeneous classrooms. *American Educational Research Journal, 34*, 365–394.

Miller, V. G. (1993). Measurement of self-perception of intuitiveness. *Western Journal of Nursing Research, 15*(5), 595–606.

Miller, G. A., Galanter, E., & Pribram, K. H. (1960). *Plans and the structure of behavior*. New York: Holt.

Montgomery, K. J., Seeherman, K. R., & Haxby, J. V. (2009). The well-tempered social brain. *Psychological Science, 20*(10), 1211–1213.

Morris, C. D., Bransford, J. D., & Franks, J. J. (1977). Level of processing versus transfer-appropriate processing. *Journal of Verbal Learning and Verbal Behavior, 16*, 519–533.

Morrison, J. E., & Fletcher, J. D. (2002). *Cognitive readiness* (IDA Paper P-3735). Alexandria, VA: Institute for Defense Analyses.

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

Niu, W., & Liu, D. (2009). Enhancing creativity: A comparison between effects of an indicative instruction "to be creative" and a more elaborate heuristic instruction on Chinese student creativity. *Psychology of Aesthetics, Creativity, and the Arts, 3*(2), 93–98.

O'Neil, H. F. (1999). Perspectives on computer-based performance assessment of problem solving: Editor's introduction. *Computers in Human Behavior, 15*, 255–268.

O'Neil, H. F. (2002). Editor's introduction: Perspectives on computer-based performance assessment of problem solving. *Computers in Human Behavior, 18*, 605–607.

O'Neil, H. F., Wang, S., Lee, C., Mulkey, J., & Baker, E. L. (2003). Assessment of teamwork skills via a teamwork questionnaire. In H. F. O'Neil & R. S. Perez (Eds.), *Technology applications in education: A learning view* (pp. 283–303). Hillsdale, NJ: Lawrence Erlbaum Associates.

Olmstead, J. A. (1992). *Battle staff integration* (IDA Paper P-2560). Alexandria, VA: Institute for Defense Analyses.

Orlansky, J., & Thorpe, J. (Eds.). (1992). *73 Easting: Lessons learned from desert storm via advanced distributed simulation technology* (IDA Document D-1110). Alexandria, VA: Institute for Defense Analyses (DTIC/NTIS No. ADA 253 991).

Perrea, J., Mohamadou, G., & Kaur, S. (2010). The use of objective structured self-assessment and peer feedback (OSSP) for learning communication skills: Evaluation using a controlled trial. *Advances in Health Sciences Education, 15*(2), 185–193.

Potosky, D., & Ramakrishna, H. V. (2002). The moderating role of updating climate perceptions in the relationship between goal-orientation, self-efficacy, and job performance. *Human Performance, 15*(3), 275–297.

Pritchard, J. S., Bizo, L. A., & Stratford, R. J. (2006). The educational impact of team-skills training: Preparing students to work in groups. *British Journal of Educational Psychology, 76*, 119–140.

Pulakos, E. D., Arad, S., Donovan, M. A., & Plamondon, K. E. (2000). Adaptability in the work place: Development of a taxonomy of adaptive performance. *Journal of Applied Psychology, 85*, 612–624.

Reed, S. K. (1987). A structure-mapping model for word problems. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 13*, 124–139.

Royer, J. M., Cisero, C. A., & Carlo, M. (1993). Techniques and procedures for assessing cognitive skills. *Review of Educational Research, 63*, 201–243.

Rubin, R. B. (1985). The validity of the communication competency assessment instrument. *Communication Monographs, 52*(2), 173–185.

Rubin, R. B., Welsh, S. A., & Buerkel, R. (1995). Performance-based assessment of high school speech instruction. *Communication Education, 44*(1), 30–39.

Salas, E., & Cannon-Bowers, J. A. (2000). The anatomy of team training. In S. Tobias & J. D. Fletcher (Eds.), *Training and retraining: A handbook for business, industry, government, and the military* (pp. 312–335). New York, NY: Macmillan.

Salas, E., Rozell, D., Mullen, B., & Driskell, J. E. (1999). The effect of team building on performance: An integration. *Small Group Research, 30*, 309–329.

Salmon, P. M., Stanton, N. A., Walker, G. H., Jenkins, D., Ladva, D., Rafferty, L., et al. (2009). Measuring situation awareness in complex systems: Comparison of measures study. *International Journal of Industrial Ergonomics, 39*(3), 490–500.

Salomon, G., & Perkins, D. (1989). Rocky roads to transfer: Rethinking mechanisms of a neglected phenomenon. *Educational Psychologist, 24*, 113–142.

Saus, E. R., Johnsen, B. H., Eid, J., Riisem, P. K., Andersen, R., & Thayer, J. F. (2006). The effect of brief situational awareness training in a police shooting simulator: An experimental study. *Military Psychology, 18*(Suppl), S3–S21.

Schraw, G., & Denison, R. S. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology, 19*, 460–475.

Schroeder, J. E., Dyer, F. N., Czerny, P., Youngling, E. W., & Gillotti, D. P. (1986). *Videodisc interpersonal skills training and assessment (VISTA): Overview and findings, Volume 1* (ARI Technical Report 703). Fort Benning, GA: U S Army Research Institute Field Unit.

Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological Review, 63*, 129–138.

Simon, H. A. (1992). What is an explanation of behavior? *Psychological Science, 3*, 150–161.

Sloutsky, V. M., & Fisher, A. V. (2008). Attentional learning and flexible induction: How mundane mechanisms give rise to smart behaviors. *Child Development, 79*(3), 639–651.

Slovic, P., Lichtenstein, S., & Fischoff, B. (1988). Decision making. In R. C. Atkinson, R. J. Herrnstein, G. Lindzey, & R. D. Luce (Eds.), *Stevens' handbook of experimental psychology, Volume 2, Learning and cognition* (2nd ed., pp. 673–738). New York: Wiley.

Sobocan, J., & Groarke, L. (Eds.). (2009). *Critical thinking education and assessment: Can higher order thinking be tested?* London, ON: Althouse.

Soest, T., Mossige, S., Stefansen, K., & Hjemdal, O. (2010). A validation study of the resilience scale for adolescents. *Journal of Psychopathology and Behavioral Assessment, 32*(2), 215–225.

Soliman, A. M., & Mathna, E. K. (2010). Metacognitive strategy training improves driving situation awareness. *Social Behavior and Personality, 37*(9), 1160–1170.

Spies, R. A., Carlson, J. A., & Geisinger, K. F. (Eds.). (2010). *Eighteenth mental measurements yearbook*. Lincoln, NE: University of Nebraska Press.

Sternberg, R. J. (1997). *Successful intelligence*. New York: Plume Books.

Sternberg, R. J. (2006a). Creating a vision of creativity: The first 25 years. *Psychology of Aesthetics, Creativity, and the Arts, 5*(1), 2–12.

Sternberg, R. J. (2006b). The Rainbow Project: Enhancing the SAT through assessments of analytical, practical, and creative skills. *Intelligence, 34*, 321–350.

Sternberg, R., & Hedlund, J. (2002). Practical intelligence, g, and work psychology. *Human Performance, 15*, 143–160.

Sternberg, R., Roediger, H. L., & Halpern, D. (2006). *Critical thinking in psychology*. New York: Cambridge University Press.

Thorndike, E. L. (1913). *Educational Psychology* (Vols. 1 & 2). New York: Columbia University Press.

Thorndike, E. L., & Woodworth, R. S. (1901). The influence of improvement in one mental function upon the efficiency of other functions. *Psychological Review, 8*, 247–261.

Tobias, S., & Duffy, T. D. (Eds.). (2009). *Constructivist theory applied to education: Success or failure?* New York: Taylor and Francis.

Tobias, S., & Everson, H. T. (2009). The importance of knowing what you know: A knowledge monitoring framework for studying metacognition in education. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Handbook of metacognition in education* (pp. 107–127). New York, NY: Routledge/Taylor & Francis Group.

Tobias, S., & Fletcher, J. D. (2011). *Computer games and instruction*. Scottsdale, AZ: Information Age Publishing.

Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review, 80*, 352–373.

Tusaie, K., & Dyer, J. (2004, January–February). Resilience: A historical review of the construct. *Holistic Nurse Practice*, 3–8.

Urban, K. K. (2004). Assessing creativity: The test for creative thinking—drawing production (TCT-DP): Concept, application, evaluation, and international studies. *Psychology Science, 46*(3), 387–397.

Walker, H. S., & Schneider, W. (Eds.). (2009). *Metacognition: Strategy use, and instruction*. New York: Guilford Press.

Weitzul, J. B. (1992). *Evaluating interpersonal skills in the job interview: A guide for human resource professionals*. New York: Quorum Books.

Wickens, C. D., & Flack, J. M. (1988). Information processing. In E. L. Wiener & D. C. Nagel (Eds.), *Human factors in aviation* (pp. 111–155). San Diego, CA: Academic.

Zaccaro, S. J. (2009). *Principles and strategies for developing skills that promote adaptive performance*. Arlington, VA: US Army Research for the Behavioral and Social Sciences.

Chapter 3
Learning and Assessment: Twenty-First Century Skills and Cognitive Readiness

Eva L. Baker

3.1 Twenty-First Century Skills and Cognitive Readiness: An Introduction

"Twenty-first century skills" is a widely adopted metaphor for a class of cognitive skills and social and affective competencies that are thought to be essential for future education and training. They are pertinent whether the goal is to develop the individual capacity to make the student ready for further or higher education, technical training or college, or to consider the types of skills thought to be of high value in the current and future work environments, including the military. In many discussions of cognitive skills the emphasis is often on the selection of individuals with generally measured aptitudes, such as intelligence or creativity.

The problem is different from a training and education perspective. With the focus on how to change skills and proclivities, many argue that twenty-first century skills should be developed within the context of explicated content domains, such as mathematics, science, and history. They may also augment the typical definitions of skill areas, such as literacy in the national language or additional languages. A second, relatively recent idea in workplace or academic settings is that it is sufficient to teach a cognitive skill, such as problem solving for instance, in a specific domain, such as algebra. Others, including this author, believe otherwise, and propose a phased approach. First, attention must be directed to helping the student learn to apply the skill in the principal areas of content. Second, to demonstrate significant competence in the domain, performance must traverse the full breadth and sufficient depth of the content domain of interest. Third, the focus on instruction and assessment should be directed to building and verifying that the learner has acquired a set

E.L. Baker (✉)
National Center for Research on Evaluation, Standards, and Student Testing (CRESST)/UCLA, 301 GSE & IS Building, 300 Charles E. Young Drive North, Mailbox 951522, Los Angeles, CA 90095-1522, USA
e-mail: eva@ucla.edu

H.F. O'Neil et al. (eds.), *Teaching and Measuring Cognitive Readiness*,
DOI 10.1007/978-1-4614-7579-8_3,

of principles incorporated into a mental schema or pattern. This step is important if trainees are expected to retrieve efficiently the key aspects of the cognitive demand rather than to memorize surface features of a procedure (Sweller, 2003). Finally, attention is not only required for a broad range of content, but more explicitly to the notion that the attributes of the situation, constraints, elements of content, and quality of solution or action may in fact change simultaneously but to different degrees. The ability to respond to such unpredictable new states is the fundamental difference in the use of the term "cognitive readiness" as opposed to twenty-first century skills. The expectation is that "readiness" implies the ability to undertake an unforeseen assignment. To achieve such readiness, the trainee or learner must be exposed to a sufficiently diverse set of conditions, situations, and problem settings so that their ability to transfer their learning to a new setting will be developed (relative to applying the schema described above).

In education, to date, most assessments do not explicitly call out twenty-first century or less modern formulations of cognitive demands, such as adaptation, risk taking, or situation awareness. They instead over-cue on the content knowledge itself, applied in routine settings. Even in research (Baker, 1997, 2007a, 2007b) the focus has been on designing learning and assessment tasks in the context of tried and true domains. This is true even when the goal for content knowledge has been improved conceptual refinement or greater specificity (see Common Core State Standards; National Governors Association Center for Best Practices, Council of Chief State School Officers, 2010). The continued explosion in knowledge suggests that routine contexts or notions of fixed domains are obsolete. Learners will need twenty-first century skills or else must relearn all the contexts and content to which these skills may apply, clearly an impossible goal. If adaptability to an unforeseen future is important, and analysts predict new careers and tasks within 5 years that are largely unknown today, learners must be comfortable with applying principles and schema they have learned as well and to determine how such schema may be modified in order to meet new requirements. Therefore, in addition to schema and transfer performance, the learner needs to develop unforeseen new skill sets which may be combinations or modifications of methods taught to solve problems, reason, or make decisions.

To summarize, at this point, in military training and in the world of work, the emphasis is on the application of some of the twenty-first century skills, concentrating on the variable context of emerging and uncertain situations. In educational research on academic learning, the rendition of "uncertainty" of future context has been often limited and focused on "transfer" situations, that is, tasks where the learner needs to apply the skill to a heretofore unlearned situation, domain, or explicit set of constraints. The difference between "academic" settings and military or workplace training may be in a figure-ground difference, where emphasis on content, skill level, or context is a matter of perspective, but different perceptions have importance for the design of both learning and assessment systems. That is, whether one primarily sees skills embedded in content or whether one's attention is on adapting to changing contexts will modify one's approach to design of assessments, and of the learning experiences that precede them. This difference in perspective may be one useful marker for deciding whether one is in the twenty-first century skills or cognitive readiness basket.

3.2 A Specification of Twenty-First Century Skills and Cognitive Readiness: A Partial List

What might be the range of skills included in twenty-first century skill sets and what is their relationship to older cognitive skills and the construct of cognitive readiness? Any list of them is not purely original, for they overlap in common language terms, sometimes using related synonyms to mean similar ideas. Further, any list would be a partial one given the state of the art. I have collapsed twenty-first century skills into three major categories: (1) intellectual cognitive processes, (2) socially oriented processes, and (3) intrapersonal skills. The first and third focus on personal development skills, either those that are outward-looking, such as how do I scan to identify key features of a problem, or inward-directed, for instance, how do I manage my own cognitive (or affective) processes to achieve my desired goals.

For the most part, such tasks in schools and training are embedded in particular subject matter. A child may be asked to figure out how to cross a river, using given objects and applying fundamental principles of force and motion, including momentum and friction. A high school student might be asked to determine the optimum shape of a figure that meets given or inferred sets of constraints. A naval officer might need to interpret signals and signs to determine whether to launch a defensive action. An internal medicine doctor might need to determine a diagnosis for muscle pain and suggest a course of action for the patient.

3.2.1 Intellectual Cognitive Processes

In the area of intellectual processes are those cognitive tasks that require sets of related subprocesses to achieve a particular goal. Consider problem solving (Baker & Mayer, 1999) which normally involves a different set of thinking skills under conditions where the problem is clearly defined as opposed to a situation where the problem is not well specified (Feltovich, Spiro, Coulson, & Feltovich, 1996). After the identification and verification of the problem, the learner or examinee must represent the problem clearly, verbally, graphically, or symbolically and determine whether the task instructions demand a "correct" or convergent answer or if they permit a divergent response requiring a new combination of learned skills, or an innovative strategy. If convergent solutions are called for, then the learner must find the optimal path and match the designer's general solution. Procedures to verify that the posited solution solves the problem may be related to another task.

Some of these skills have a long and relatively unchanging set of definitions, for instance the difference between inductive reasoning (from examples) and deductive approaches (reasoning from a premise). Work in this area goes back to classical times, but a more recent publication by Johnson-Laird (2006) summarizes evidence in this area. A newer principle may be the cognitive skill of search, where students are encouraged to conduct searches, but to use a model of evidence to validate their findings. An earlier literature, mostly in the field of information science, has given way to the ability to use search engines well, particular tips, for instance, offered at almost every university library, and varying sets of criteria to judge quality of output (see Dragan, Tepper, & Misso, 2011, for an example). The ability to judge criteria of quality will be increasingly needed as technology repositories may either be crowd-sourced (e.g., Wikipedia) or not subjected to any quality control. Another area of interest is the understanding of declarative, procedural, and systemic knowledge. This understanding is critical for as noted earlier, cognitive skills are embedded in subject matter and situations. Sources for the tripartite definitions of understanding (What, How, and Why correspondent to declarative, procedural, and systemic knowledge) have a long history. Writings in this area include a series of articles by Alexander and Judy (1988) reviewing the literature in the area, and an excellent earlier piece by de Jong and Ferguson-Hessler (1996) on types and qualities of knowledge. These categories have been combined with Bloom's *Taxonomy of Educational Objectives* (1956) in a revision by Anderson and Krathwohl (2001).

A list of potential twenty-first century skills focused on intellectual cognitive processes is listed below. When these are engaged effectively, "on the fly" without much preparation, they merge into the realm of cognitive readiness.

1. Adaptive problem solving
2. Decision making
3. Situation awareness
4. Reasoning, inductive and deductive
5. Searching for missing resources
6. Understanding declarative, procedural, and systemic knowledge in particular subdomains

Two points are obvious. First, each of the skills of intellectual cognitive processes does not operate in isolation. A task nominally focused on one skill may depend upon the application of others. For example, problem solving may require situation awareness (Koenig, Lee, Iseli, & Wainess, 2009) and decision making (Lee, Bewley, Jones, Min, & Kang, 2009; Lee, Jones, & Min, 2009). Second, each of the skills may take many different surface forms. These skills play out differently in domains composed of different principles, concepts, facts, and routine approaches. Going back to the problem-solving task area above to illustrate, consider an ill-defined problem. A problem may be ill-defined in a number of ways. It may be vague without sufficient information to guide strategies for a solution. The respondent will then need to generate alternative, plausible interpretations of the problem, represent them in an appropriate fashion, and then proceed to apply strategies or procedures for solutions. It is clear that there may be loops and returns to the beginning if the problem identified is not what was intended by the designer.

A problem might also be ill-defined because part of it is hidden or occluded. That is, the learner may need to wade through extraneous material to find the real problem, stimuli that might be verbal, visual, or both. The task writer may work very hard at deception, trying to lead the respondent to a wrong interpretation. To get through this type of ill-defined problem, the learner must know enough about the task situation and the content domain to avoid false lures. If in the problem setup, the designer has included explicit or less obvious constraints, then the respondent must use skills of situation awareness to be able to focus on the real problem. The learner will also need to use reasoning skills, if he or she is to discard less central material and determine which variables are critical determinants of the task. At key points of a complex, multistep problem, the respondent will need to make decisions. These decisions will be made on the basis of prior experience, with knowledge provided or obtained relevant to the task, and by the application of reasoning.

When analysts consider such skills as a "trait" or innate individual differences, they are positing that a good "problem solver," logical thinker, or detector of change in environmental situations will be able to apply these skills generally, without specific training in or across domains. In other words, the cognitive skills are used independently of the particular content domain with about the same level of skill. It turns out, however, that individuals may have their abilities bounded by related domains, for instance, good problem solving in math will not bleed over to the same respondent's behavior in literature. What we are positing, however, is that training that involves transfer of skills first within a broad domain, then in varying situations applicable to the domain and the intellectual skill, will need to be augmented by training that requires application across content domains. Whether such a domain-independent set of performance skills can be developed remains questionable. Each domain typically requires relatively deep declarative, procedural, and systemic knowledge before the respondent can use strategies to solve problems or make decisions. It is somewhat unlikely that ordinary rather than extraordinary minds will acquire deep knowledge over a wide range of domains. Thus, the plausibility of the domain-independent application of intellectual cognitive skills depends on the limits of time, interest, and capacity to learn domains that are far afield from one another.

3.2.2 Social and Interpersonal Skills

A second class of twenty-first century skills involves a set of socially oriented skills that include cognition, but require its application in interpersonal situations. These competencies may involve the collaborative nature of work on the one hand, or the ever present need to communicate to obtain approvals for plans, to discuss and understand work, and to report it to a range of audiences. Moreover, socially oriented skills do not heavily depend upon sunny or outgoing personalities. For instance, the area of collaboration requires the learner to be able to clarify goals of the team or collaboration, and to modify behavior to acknowledge the value of ideas, even in situations where they may differ in opinion with those in the group.

We provide a partial list of such skills below:

1. Teamwork
2. Collaboration
3. Help

4. Social situational awareness
5. Communication—productive and receptive

To succeed at work using the intellectual skills, the learner may need to depend upon other people as a resource on the one hand, or provide only part of the effort needed for a solution that will involve many players. Writers about teamwork (Salas & Cannon-Bowers, 2001) have created a set of component skills that may involve providing leadership, feedback, motivation, redirection, clarification, incentives, and effort in order to succeed in the task. Not all tasks call for all of these components, nor should it be imagined that any single person is required to provide all of them for a given task.

But from a learning perspective, to be a good team or group member, some of them will be required. For example, to be a leader requires from time to time different tacks for different requirements. Clarity of goals may be a constant requirement, but it may be wise to have a group clarify goals rather than have the goal specified always by a particular individual. Obviously, the manner of application depends upon organizational hierarchy and individual status. To make the right choice, the team member must be aware of the various states of the group or team, including their willingness to participate, their role, and other information, such as their experience or desire to try new things. Research on the topic of the social situational awareness is often called empathy (see Keltner, 2004). Here it is differentiated because it is formulated as a skill to be learned as opposed to an attribute that is inherent.

It is clear that interactions in the social realm, whether face-to-face, continuously or occasionally, through media, or transmitted by text, depend upon some level of expertise in communication. The skills are required to formulate, compose, and explain important tasks or to ask and answer key questions. Another aspect of communication is the sensitivity to the use of appropriate language, suitable to the audience, the organization, task, and specific interpersonal contexts, including, where relevant, cultural issues. These then form an evolving set of related skills pertinent to the domain of social situational awareness, combining skills that are both intellectual and interpersonal.

As interpersonal skills may have strong experiential, cultural, and personality bases, it is unclear that they can, in totality, be trained or taught. However, key components, for instance communication, have a long history of being learned and assessed. Some studies look at the relationship of receptive (listening) or reading and productive (speaking or writing) at very specific levels (see for instance, Guess, 1969) or as a more specific set of interpersonal skills (Hargie, 2006). Vygotsky (1978) saw social interaction and communication at the heart of higher psychological processing. Elements of collaboration or teamwork, especially understanding and executing specific steps or roles related to identify tasks, also can be well taught. In working on teams, the interpersonal components will develop over time, over stress conditions, and in varied task situations. Tasks vary as those that can be completed independently by members of a group or where each team member makes a unique and interdependent contribution to the attainment of the goal (see Webb, 1985).

3.2.3 Intrapersonal Skills

Intrapersonal skills are those personal behaviors and internal thought processes that can be systematically acquired or enhanced by instruction or learning experiences and for which evidence of change can be directly or indirectly inferred. A partial list of such skill sets follows.

1. Planning
2. Self-monitoring
3. Emotional control
4. Self-control
5. Risk taking
6. Mental effort
7. Attributions of success and failure
8. Self-efficacy
9. Resilience

Whereas some of these skills may be clustered under categories of metacognition, for instance planning and self-monitoring, others have a more emotional component. For example, individuals can learn to control emotions, to practice a more balanced personal demeanor through mind-body regimens, or to estimate risk and change their propensity for risk taking related to situations, e.g., costs, likely success, consequences of failure. (There is general evidence that each of these can be taught to some degree.) These self-controlled skills can be learned, and may involve attribution (Weiner, Graham, & Reyna, 1997), self-efficacy, and resilience. These are components of self-awareness. The process of emotional control, under stress, derives from personality psychology (see Roger & Nesshoever, 1987, for an example of measurement). There are many examples with long research histories, for example, desensitization research, that ranges from those documenting therapeutic approaches (Paul & Bernstein, 1976), effects of sexism (Linz, Donnerstein, & Penrod, 1988), or the effects of violent media and games (Anderson et al., 2010; Cline, Croft, & Courrier, 1973).

Risk taking, moderated by a sense of payoff, is another intrapersonal skill that has more currency in the context of self-motivated actions, leadership, and entrepreneurship. This work has a lengthy history as well (see for example, Brockhaus, 1980) and has been investigated often in the context of management schools.

The interacting areas of effort and attribution have been well summarized by Graham (1991) and Perry, Stupnisky, Hall, Chipperfield, and Weiner (2010), where researchers have first shown a relationship between success and attributions to personal effort, and have developed a training regimen to develop such concepts in those who think that they have been either selected to fail or have no control over their learning. This line of inquiry is singularly related to studies of stereotype threat and how to overcome such perceptions by African Americans (Logel, Walton, Spencer, Peach, & Mark, 2012). Claude Steele (Steele, Spencer, & Lynch, 1993) has continued to probe the role that self-perception plays to support resilience under socially threatening conditions.

Intrapersonal processes may or may not be as amenable to instruction, depending upon how the construct is formulated, and the cultural context, age, and experience of the learner. In this twenty-first century skill discussion, however, these elements are imagined to be changeable through education, training, and experience. Individual differences apply to each of these areas, and intensive training may enhance intrapersonal skills to a higher level. On the other hand, some background differences, for instance, experience with failure, may make it harder for training to be effective.

3.3 Cognitive Readiness Revisited

If the elements exemplified in the intellectual, social, and intrapersonal twenty-first century skills are meant to be used in educational and training situations, what is the difference between them and the elements of cognitive readiness described by Fletcher (2004), Fletcher and Wind (this volume), and O'Neil, Lang, Perez, Escalante, and Fox (this volume). At some level, the cognitive readiness notion emphasizes the ability to be agile, adaptive, and prepared for uncertainty. While many elements of twenty-first century skills also share that emphasis, it is fair to say the term "readiness" means the ability to act in unpredictable situations. For the most part, twenty-first century skills are vested in institutionally based learning, schools, university, or world of work. When the term "cognitive readiness" is used in the military context, it engenders an image of rapidly evolving situations, unforeseen challenge and constraints, and the requirement of rapid rather than reflective or long-term analysis.

3.3.1 Learning and Assessment Options for Cognitive Readiness

These differences in specificity or boundaries imply important alternatives in the design of situations intended to develop or to measure performance. In cognitive readiness, the emphasis is on the *combination* of individual differences and trained or learned skills. The "readiness" depends upon the interaction between both sources of expertise of the individual and from the given requirements of a situation. In designing learning experiences or exercises to lead to these particular outcomes, alternative combinations of instructional strategies are needed. For example, if the goal were to teach a broad-based construct such as agility of thinking under pressure, then the training for such an approach might be based on a wide range of different, unexpected situations where the respondent either needed to draw from experience or invent a solution within a tight time limit.

If viewed as an individual difference variable, the measurement of such ability would suggest the use of more classical test theories that depend upon normal distributions of the trait. In such measure development, the construct is hypothesized, items or tasks in a diverging set of situations are generated, and the item set is culled, based upon the degree the responses follow predicted patterns. Do they cluster together in logical ways, can they be compared to other known measures, and do they result in discriminations among individuals with more or less of the trait or aptitude? Practically speaking, the use of existing measures for some of the areas, such as creativity, may be desirable if combined in a test battery that generates a profile of the individual. Such a process, if undertaken as a test development procedure, would require many responses to determine the operational definition of the construct, and then validation either under the circumstances of use (that is, surprising, threatening situations). Alternatively, scores and profiles of those thought to possess the trait could be contrasted to those who were not thought likely to have the trait. Although the trait itself may be normally distributed, with some people having a lot, and some having little, the next step in the process would be to determine how much effect feasible amounts of training might have in either compensating for low levels of trait possession or increasing the expertise of those in the middle range to a high level of effectiveness. This way of thinking about the cognitive readiness notion is looking trait by trait at the measured abilities of individuals, and then placing them in criterion situations that nominally require the application of the trait. Although it might be academically clean to conduct such studies in a step-wise way, it is more likely that correlated traits will be co-evaluated, rather than singly established.

The challenge in cognitive readiness is not so much in the tasks to measure traits or aptitudes, but rather in setting up the validation situations to represent a range of experiences demanding cognitive readiness. The definitional challenge of unpredictable situations (unpredictable to whom?), and the enactment of a validation sequence with sufficient breadth to allow confidence in performance is a far more daunting task than first meets the eye.

3.4 The Role of Ontologies to Design Content to Combine with Twenty-First Century Skills or Cognitive Readiness Constructs

The second step in implementing any of the three types of twenty-first century skills involves the use of an ontology, or map of the content domain(s) of interest. An ontology is a graphic representation of language with the following characteristics: a network of nodes, a set of links describing the relationship among nodes, and a database which will modify the direction or arrangement of nodes and links based on performance information. If an ontology is a graphical representation of a content domain to be used for learning or assessment, what is its construed properties? An ontology features principles, concepts, key knowledge, and procedures. These are depicted as nodes in a network. The links among nodes have direction and meaning. They may convey, in a hierarchical representation, the components that are subsets of others, and range from desired higher order complex content all the way down to fundamental principles and facts that beginners are expected to know. Mathematics is an excellent area to display hierarchical strands of topical domains.

For different content, the ontology may also be structured in an appropriate, active way, where principles, concepts, and facts are linked in the ontology to convey chronology, themes, or principles. Political history is an example of a domain that may have a structure based on chronology. Another different structural form of a content ontology may illustrate the relationships among nodes (containing principles, concepts, or examples) in terms of their mutual or directional influence on one another. The strongest case of such links might be those that exemplify causal relations. In most fields of study, one will find a mix of relationships, some reflecting part-whole relationships, others more thematic, chronological, or causal. In some areas, the structure may be loose where the broad domain is considered, for example, in literature, where relationships among forms, such as the novel or poetry, may be parallel or horizontal rather than hierarchical. Yet, within a literary form, for instance, stronger vertical structure may be found. In cases of plot or character development, highly explicit relationships can be developed.

The structure of the ontologies must follow the essential character of the domain. These may vary by the extent to which interpretive processes are present as opposed to specific methodologies that are intended to yield relatively clear outcomes. These domains will also differ in the light of the extent to which they represent abstract principles, addressed theoretically or empirically, for instance in physics, as opposed to domains in which each example may only be loosely joined to the next, for instance, examples of lyric poetry. Figure 3.1 depicts an ontology in algebra.

3.4.1 Common Elements of Ontologies

Independent of structure, any ontology has particular features. First it is a graphical representation. Second, there is no restriction about how many nodes may be linked to any other. In point of fact, centrality of content can be determined simply by looking at the number of nodes with the highest frequency of links. One can also determine which content is remote, and potentially nonessential, because it is less well connected to more central ideas.

This operational depiction of importance can lead to direct inferences about the design of learning systems. While some may view an ontology to be a sequence of instruction, or the optimal arrangement of a computer adaptive test, no such inference should directly be made. For instance, it is fashionable to use the term "learning progression" as if one truly knew which sequence was optimal. However, only empirical study can generate stronger hypotheses about the order of learning or of assessment for learners and settings. The type of empirical study could be undertaken to contrast the processes followed by individuals known to be competent in the domain with those of novices, or with individuals only partially taught, or able to demonstrate middling levels of performance. Other empirical work can take hypothesized arrangements of tasks and delete systematically alternative elements in order to determine which content and processes are essential to the achievement of desired goals.

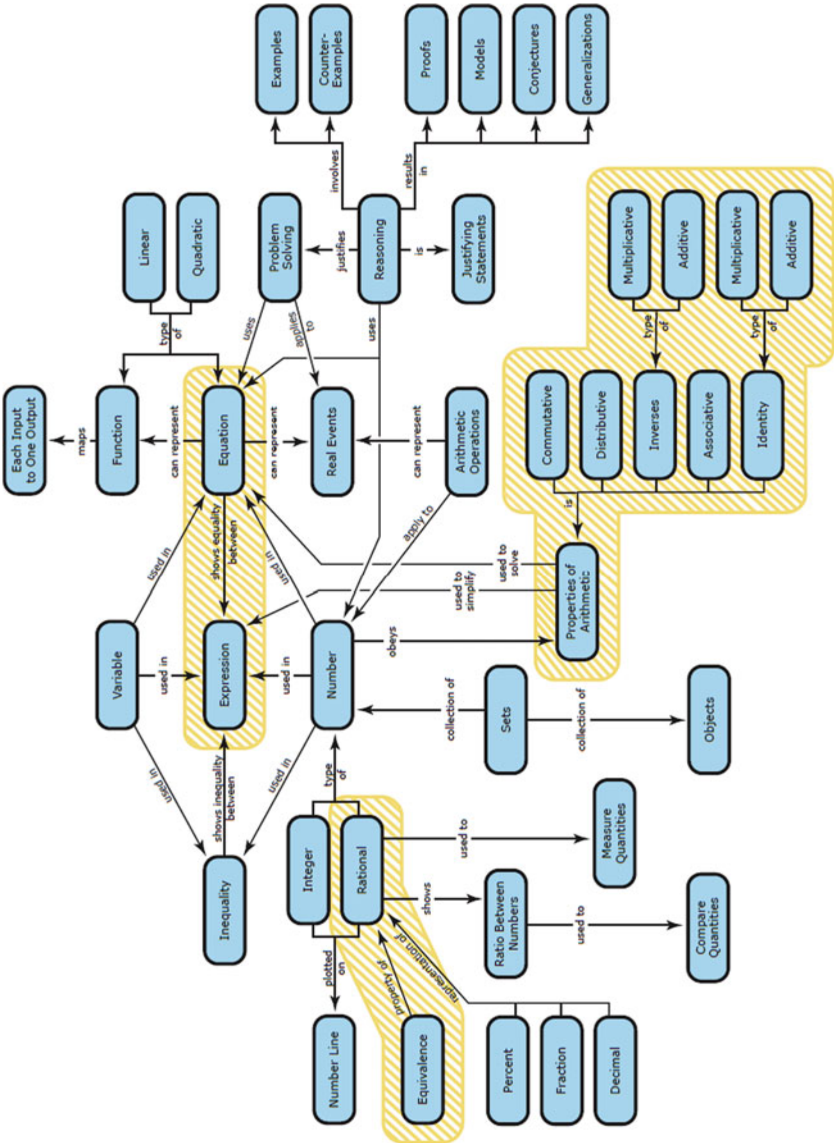


Fig. 3.1 Algebra ontology content

3.4.2 How Is an Ontology Made?

Ontologies are usually made iteratively and made by people. Imagine that an expert in a subject matter area were asked to represent the domain of interest in terms of its important topics and their relationships to other concepts, principles, or facts. One way of thinking about an ontology is that we are asking the expert to externalize and to depict symbolically a mental model of the domain. In practical ontology development, experts are carefully selected and asked to create an ontology individually (using software, e.g., the Knowledge Mapper developed by Chung & Baker, 1997, where the expert can create the structure of nodes and links using a pull-down menu of options). When multiple experts each develop an ontology, they can be superimposed graphically to show areas of agreement and differences. A typical process then requires face-to-face discussions, explanations, and reconciliation of differences by the experts, iteratively, until consensus is reached.

In addition, the recent research undertaking by Chung, Niemi, and Bewley (2003) and more recently by Iseli (2011) indicates that ontology development can begin or be augmented in process with the analysis of documents relevant to the domain. Automated extraction of key ideas in referent texts, articles, and other documents can be achieved using natural language processing placed in a network representation and included in the material that the experts are to reconcile. Such content ontologies are in development for mathematics from kindergarten through secondary schools (Iseli, 2011; Iseli, Koenig, Lee, & Wainess, 2010), history (Phelan, Dai, Valderrama, & Herman, 2011), biological science (Phelan et al., 2011), and language arts (Phelan et al., 2011).

3.4.3 Blending Twenty-First Century Skills and Ontologies

In the ideal case, ontologies of twenty-first century skills should be merged with a content ontology to create an integrated architecture to guide learning and assessment. At CRESST, we have made some progress in creating ontologies for twenty-first century skills, in the areas of adaptive problem solving (Mayer, 2010), communication (Phelan et al., 2011), situation awareness (Koenig et al., 2009), and teamwork (O'Neil, Wang, Lee, Mulkey, & Baker, 2003). Each of these ontologies is based on theoretical and empirical analyses of the process domains. For instance, in Mayer's problem-solving ontology, nationally recognized cognitive psychologists were used as experts and asked to create an ontology, one that is still under revision to refine its content and structure. In the skill of situation awareness, research by Endsley (1995) and her colleagues (Bolstad, Endsley, & Cuevas, this volume) was essential, and in the teamwork area, frameworks and empirical studies by Salas and Cannon-Bowers (2001), and O'Neil and colleagues (O'Neil, Chuang, & Chung, 2003; O'Neil, Chung, & Brown, 1997; O'Neil, Wang, Chung, & Herl, 2000). It is our intention to continue to document the process by which intellectual and social skill ontologies are developed and combined with content domain ontology for the purposes of assessment, simulation, and game design (Bewley, Lee, Jones, & Cai, this volume).

Cognitive readiness, however, may seem antithetical to an approach that uses an ontology as the way to represent “all possible” content. The unpredictability that uniquely defines cognitive readiness is not in the details of the content, although one might posit that less frequent or more unusual content has a place in the definition of “unexpected.” The unpredictability may be more likely related to the unusual situation in which an individual is placed. There, the benefit of well-learned knowledge structures is in the rapid retrieval and testing relevancy against the situational requirements. There is not much evidence that this is the process that leads to achievement, survival, and avoidance of error, but there is a clear research path that could be taken to determine the components of knowledge needed to be adaptive in unpredicted environments.

3.5 Model-Based Learning and Assessment of Twenty-First Century Skills Embedded in Content and Cognitive Readiness in Unusual Situations

Design using twenty-first century skills, cognitive readiness, and content ontologies must be realized in assessment and learning situations and systems. Let us focus on assessment. We have developed a model for the development of assessments and its history has evolved since 1991 (Baker, 2007c; Baker, Chung, & Delacruz, 2012; Baker, Freeman, & Clayton, 1991). In the model of assessment we begin not with content specifications, which is the usual practice, but with the selection of twenty-first century skills that will be embedded in the content domain of the ontology measure. Reasons for this explicit inclusion of relevant twenty-first century skills can be explicated. First, it is done to assure that the intellectual depth of processing is included on the measure as intended by statements on standards or doctrine. If not made explicit, many tests are found to overemphasize recognition or repetitive procedures simply because those are easy to generate. Second, the operational definitions of twenty-first century skills enable the determination of which relevant content should be given higher weight in the intellectual skill domain. Third, particular intellectual skills suggest assessment formats to optimize caliber of measurement. For example, adaptive problem-solving demands that the respondent create an original answer, whereas a problem identification measure might combine selected responses, for example, which is the best statement or representation of the problem, with a verbal explanation about why the choice of problem statement was made.

An explicit cognitive model of a twenty-first century skill allows the generation of one or more templates or sets of modular objects that can be used in combination to build assessments. This modular feature looks forward to partially automated assessment design using computer-based authoring systems. Next, the use of a common model task for assessment will increase the likelihood of coherent sampling within a domain. Construct-irrelevant task features (Messick, 1989) or item types that add noise to the understanding of student performance will be identified and reduced. This focus on measuring the desired and relevant outcomes will affect positively reliability of findings and help detect real change as a function of intervention or experience. Moreover, the use of models will permit the development of subsequent extensions of the measures at a lower cost, because there will be three sources of guidance: the twenty-first century skills, the ontology of relevant content, and the assessment task model. This economic utility should not be underestimated. The templates can be reused and different situations, content, or responses can be inserted, which will support longitudinal interpretations of growth. Moreover, in the case of open-ended responses, scoring criteria or rubrics can also be reused and will greatly reduce cost. Furthermore, if rubrics are at a high level of connection to the content ontology and twenty-first century skills ontology, teachers and students can be recipients of more transparent requirements for learning and assessment.

3.6 Validity of Twenty-First Century Skills and Cognitive Readiness Purposes

The approach to validity and other relevant instances of technical quality should be explicit. Validity is purpose-driven. Validity is a chain of inferences linking the purpose of the assessment to data and subsequent inferences about the quality of decision the assessment yields (see American Educational Research Association, American Psychological Association, & National Council on Measurement in Education, 1999; Messick, 1989). With different purposes, types of data and relevant inferences will vary. If the major use of data is formative assessment, for instance to help the designer decide how a simulation or game should progress based on learners' prior performance, then the data of interest will be relatively granular, and focused on misconceptions, hesitations, and errors that need to be addressed in a revision of the system or perhaps just relevant to the alternative paths provided for individual students. If the data are to be used to assess the acquisition of a set of skills for certification or accountability, on the other hand, then the validity evidence would consist of the determination whether those skills were also represented in the repertoires of the skilled performers or other individuals who had nominally achieved the desired goals.

In addition, if public reporting for accountability were of interest, for instance, to categorize schools in terms of their own effectiveness, then information about learner, student, and teacher groups; mobility amount and types of instruction; and different levels of satisfactory achievement would be considered to infer validity. In all cases, however, data management systems would be needed to report outcomes to designers, teachers, or policymakers. They would need to be organized so that progress along skill dimensions could be tagged, as well as other potential variables such as content, situational, and linguistic complexity. This database function refers back to the larger use of ontologies, recalling that they serve not only a graphical representation role, but guide the design of ongoing data management systems and reporting.

3.7 Summary

This chapter has provided perspectives on twenty-first century skills and cognitive readiness, justifying our interest in them in the face of increasing uncertainty. The set was divided in three ways: (1) intellectual cognitive skills; (2) socially oriented skills, and (3) intrapersonal skills. Comments related to selection or training of these skills were followed by an analysis of the term "cognitive readiness" and its intent in explicitly dealing with details of transfer and preparation to confront unexpected requirements. The role of content domains with respect to twenty-first century skills was treated, and a method for representing the details of content and twenty-first century skills was developed. The graphical approach to representing content was defined and described in the concept of ontology or a symbolic representation of content, relationships, and structure. Uses, experiences, and research options were discussed with regard to ontologies as a method for identifying important skills for assessment sampling and for learning design. A brief acknowledgment of the need for ontologies for selected twenty-first century skills was described. The approach to creating assessments based on twenty-first century skills and content ontologies, "model-based assessment" (Baker, 1997, 2007b) was described in terms of its utility in creating higher level tasks, increasing technical quality of measures, and reducing cost.

A brief discussion of validity included the key notions of minimizing construct-irrelevant variance and drawing correct inferences related to purpose. Finally, an extension of the database development of ontologies was presented. It is a more complex formulation of the relationship of individuals to explicit aspects of learning: skills, content, linguistics, assessment formats, experiences, and developmental trajectories. This strategy may have relevance as more automation of student experiences is systematically accomplished.

Acknowledgments The work reported herein was supported under the Office of Naval Research, PR/Award Number N00014-09-C-0813; the Defense Advanced Research Projects Agency, PR/Award Number N00014-12-C-0090; and the Educational Research and Development Centers Program, PR/Award Number R305C080015, as administered by the Institute of Education Sciences, the U.S. Department of Education. The findings and opinions expressed in this chapter do not necessarily reflect the positions or policies of the Office of Naval Research, the Defense Advanced Research Projects Agency, or the Institute of Education Sciences, the U.S. Department of Education.

References

Alexander, P. A., & Judy, J. E. (1988). The interaction of domain-specific and strategic knowledge in academic performance. *Review of Education Research, 58*(4), 375–404.

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). *Standards for educational and psychological testing*. Washington, DC: American Educational Research Association.

Anderson, C. A., Shibuya, A., Ihori, N., Swing, E. L., Bushman, B. J., Sakamoto, A., et al. (2010). Violent video game effects on aggression, empathy, and prosocial behavior in eastern and western countries: A meta-analytic review. *Psychological Bulletin, 136*(2), 151–173.

Anderson, L. W., & Krathwohl, D. R. (with Airasian, P. W., Cruikshank, K. A., Mayer, R. E., Pintrich, P. R., Raths, J., & Wittrock, M. C.). (2001). *A taxonomy for learning, teaching and assessing: A revision of Bloom's taxonomy of educational objectives*. New York: Longman.

Baker, E. L. (1997). Model-based performance assessment. *Theory Into Practice, 36*(4), 247–254.

Baker, E. L. (2007a, August/September). The end(s) of testing (2007 AERA presidential address). *Educational Researcher, 36*(6), 309–317. Retrieved October 2, 2007, from http://www.aera.net/uploadedFiles/Publications/Journals/Educational_Researcher/3606/09edr07_309-317.pdf.

Baker, E. L. (2007b). Model-based assessments to support learning and accountability: The evolution of CRESST's research on multiple-purpose measures. *Educational Assessment (Special Issue), 12*(3&4), 179–194.

Baker, E. L. (2007c). Teacher use of formal assessment in the classroom. In W. D. Hawley with D. L. Rollie (Eds.), *The keys to effective schools: Educational reform as continuous improvement* (2nd ed., pp. 67–84). Thousand Oaks, CA: Corwin.

Baker, E. L., Chung, G. K. W. K., & Delacruz, G. C. (2012). The best and future uses of assessment in games. In M. C. Mayrath, J. Clarke-Midura, D. H. Robinson, & G. Schraw (Eds.), *Technology-based assessments for 21st century skills: Theoretical and practical implications from modern research* (pp. 227–246). Charlotte, NC: Information Age.

Baker, E. L., Freeman, M., & Clayton, S. (1991). Cognitive assessment of history for large-scale testing. In M. C. Wittrock & E. L. Baker (Eds.), *Testing and cognition* (pp. 131–153). Englewood Cliffs, NJ: Prentice-Hall.

Baker, E. L., & Mayer, R. E. (1999). Computer-based assessment of problem solving. *Computers in Human Behavior, 15*(3/4), 269–282.

Bloom, B. S. (with Engelhart, M. D., Furst, E. J., Hill, W. H., & Krathwohl, D. R.). (1956). *Taxonomy of educational objectives: The classification of education goals. Handbook I: Cognitive domain*. New York: Longman.

Brockhaus, R. H., Sr. (1980). Risk taking propensity of entrepreneurs. *Academy of Management Journal, 23*(3), 509–520.

Chung, G. K. W. K., & Baker, E. L. (1997). *Year 1 technology studies: Implications for technology in assessment* (CSE Tech. Rep. No. 459). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing.

Chung, G. K. W. K., Niemi, D., & Bewley, W. L. (2003, April). *Assessment applications of ontologies*. Paper presented at the annual meeting of the American Educational Research Association, Chicago, IL.

Cline, V. B., Croft, R. G., & Courrier, S. (1973). Desensitization of children to television violence. *Journal of Personality and Social Psychology, 27*, 360–365.

de Jong, T., & Ferguson-Hessler, M. G. M. (1996). Types and qualities of knowledge. *Educational Psychologist, 31*(2), 105–113.

Dragan, I., Tepper, K., & Misso, M. (2011). Teaching evidence based medicine literature searching skills to medical students during the clinical years—A protocol for a randomized controlled trial. *BMC Medical Education, 11*, 49. Retrieved from http://www.biomedcentral.com/content/pdf/1472-6920-11-49.pdf.

Endsley, M. R. (1995). Measurement of situation awareness in dynamic systems. *Human Factors, 37*, 65–84.

Feltovich, P. J., Spiro, R. J., Coulson, R. L., & Feltovich, J. (1996). Collaboration within and among minds: Mastering complexity, individually and in groups. In T. Koschmann (Ed.), *CSCL: Theory and practice of an emerging paradigm* (pp. 25–44). Mahwah, NJ: Erlbaum.

Fletcher, J. D. (2004). *Cognitive readiness: Preparing for the unexpected*. Alexandria, VA: Institute for Defense Analyses. Retrieved from http://www.dtic.mil/cgi-bin/GetTRDoc?AD=ADA458683&Location=U2&doc=GetTRDoc.pdf

Graham, S. (1991). A review of attribution theory in achievement contexts. *Educational Psychology Review, 3*(1), 5–39.

Guess, D. (1969). A functional analysis of receptive language and productive speech: Acquisition of the plural morpheme. *Journal of Applied Behavior Analysis, 2*(1), 55–65.

Hargie, O. (2006). *The handbook of communication skills* (3rd ed.). New York: Routledge.

Iseli, M. (2011). *Ontology development: Overview and example* (Draft CRESST Whitepaper). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing.

Iseli, M. R., Koenig, A. D., Lee, J. J., & Wainess, R. (2010). *Automatic assessment of complex task performance in games and simulations* (CRESST Report 775). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Johnson-Laird, P. (2006). *How we reason*. New York: Oxford University Press.

Keltner, D. (2004). The compassionate instinct. *Greater good*. Retrieved from http://greatergood.berkeley.edu/article/item/the_compassionate_instinct/

Koenig, A. D., Lee, J., Iseli, M., & Wainess, R. (2009, November). *A conceptual framework for assessing performance in games and simulations.* Paper presented at the Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC), Orlando, FL.

Lee, J. J., Bewley, W. L., Jones, B., Min, H., & Kang, T. (2009, November). *Assessing performance in a simulated combat information center.* Paper presented at the Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC), Orlando, FL.

Lee, J. J., Jones, B., & Min, H. (2009, April). *Assessing performance in the multi-mission team trainer*. Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.

Linz, D., Donnerstein, E., & Penrod, S. (1988). Effects of long-term exposure to violent and sexually degrading depictions of women. *Journal of Personality and Social Psychology, 55*, 758–768.

Logel, C. R., Walton, G. M., Spencer, S. J., Peach, J., & Mark, Z. P. (2012). Unleashing latent ability: Implications of stereotype threat for college admissions. *Educational Psychologist, 47*(1), 42–50.

Mayer, S. (2010). *Problem solving ontology* (CRESST Whitepaper). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Messick, S. (1989). Validity. In R. Linn (Ed.), *Educational measurement* (pp. 13–103). New York: McMillan.

National Governors Association Center for Best Practices, Council of Chief State School Officers. (2010). *Common core state standards.* Washington, DC: National Governors Association Center for Best Practices, Council of Chief State School Officers. http://www.corestandards.org/the-standards

O'Neil, H. F., Chuang, S., & Chung, G. K. W. K. (2003). Issues in the computer-based assessment of collaborative problem solving. *Assessment in Education, 10*, 361–373.

O'Neil, H. F., Jr., Chung, G. K. W. K., & Brown, R. (1997). Use of networked simulations as a context to measure team competencies. In H. F. O'Neil Jr. (Ed.), *Workforce readiness: Competencies and assessment* (pp. 411–452). Mahwah, NJ: Erlbaum.

O'Neil, H. F., Wang, S.-L., Chung, G. K. W. K., & Herl, H. E. (2000). Assessment of teamwork skills using computer-based teamwork simulations. In H. F. O'Neil & D. H. Andrews (Eds.), *Aircrew training and assessment* (pp. 245–276). Mahwah, NJ: Erlbaum.

O'Neil, H. F., Jr., Wang, S., Lee, C., Mulkey, J., & Baker, E. L. (2003). Assessment of teamwork skills via a teamwork questionnaire. In H. F. O'Neil Jr. & R. S. Perez (Eds.), *Technology applications in education: A learning view* (pp. 283–303). Mahwah, NJ: Erlbaum.

Paul, G. L., & Bernstein, D. A. (1976). Anxiety and clinical problems: Systematic desensitization and related techniques. In J. T. Spence, R. C. Carson, & J. W. Thibaut (Eds.), *Behavioral approaches to therapy*. Morristown, NJ: General Learning Press.

Perry, R. P., Stupnisky, R. H., Hall, N. C., Chipperfield, J. G., & Weiner, B. (2010). Bad starts and better finishes: Attributional retraining and initial performance in competitive achievement settings. *Journal of Social and Clinical Psychology, 29*(6), 668–700.

Phelan, J., Dai, Y., Valderrama, M., & Herman, J. (2011). *Development of learning-based assessments in literacy: Towards the goal of college readiness* (Whitepaper submitted to 2012 AERA). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Roger, D., & Nesshoever, W. (1987). The construction and preliminary validation of a scale for measuring emotional control. *Personality and Individual Differences, 8*(4), 527–534.

Salas, E., & Cannon-Bowers, J. A. (2001). The science of training: A decade of progress. *Annual Review of Psychology, 52*, 471–499.

Steele, C. M., Spencer, S. J., & Lynch, M. (1993). Self-image resilience and dissonance: The role of affirmational resources. *Journal of Personality and Social Psychology, 64*(6), 885–896.

Sweller, J. (2003). Evolution of human cognitive architecture. In B. Ross (Ed.), *The psychology of learning and motivation* (Vol. 43, pp. 215–266). San Diego, CA: Academic.

Vygotsky, L. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Webb, N. M. (1985). Student interaction and learning in small groups: A research summary. In R. Slavin, S. Sharan, S. Kagan, R. H. Lazarowitz, C. Webb, & R. Schmuck (Eds.), *Learning to cooperate, cooperating to learn* (pp. 147–172). New York: Plenum.

Weiner, B., Graham, S., & Reyna, C. (1997). An attributional examination of retributive versus utilitarian philosophies of punishment. *Social Justice Research, 10*, 431–452.

Chapter 4
A Model for Instruction and Assessment of Cognitive Readiness

Robert J. Sternberg

"Teaching and measuring cognitive readiness" seems to imply something more than just the standard methods of teaching and assessment. The question is—what more?

Readiness means first, we believe, the ability to cope with relative novelty—to handle new situations that are not only different, but different in kind from past situations. The relevant skill is creative thinking. This is not merely the kind of thinking involved in dreaming up unusual uses of a paper clip, but rather, the ability to redefine problems and see situations in new ways. The second relevant skill is analytical thinking: One needs to evaluate whether one's creative response is a good one. The third relevant skill is practical thinking: One needs to ensure that one's response is practical with respect to the constraints of the situation. And the final skill considered here is wisdom, meaning that one must ensure that the response ethically serves the common good and not just one's own.

Consider two examples. The first is financial and the second is political.

First: John is trying to grow the finances of his organization, which have been battered by a recession. He comes up with an idea for an alternative investment. It may make the organization a lot of money. The idea is novel—it is creative. But is it good? He analyzes the investment, doing a cost-benefit analysis. It looks good. Then he asks whether the investment is practical: Will it make sense in the context of the organization's overall investment strategy, and as important, will he be able to persuade the investment committee to make the investment? Finally, he asks himself whether the investment is ethical and wise. Has he used an insider trading information? No—he is ok on that. Will the investment be sound in the long term as well as the short term? John knows that many investments that look good in the short run prove to be horrible mistakes in the long run. Is the investment in a company that

R.J. Sternberg (✉)
Office of Academic Affairs, Oklahoma State University, Whitehurst 101,
Stillwater, OK 74078, USA
e-mail: Robert.sternberg@okstate.edu

H.F. O'Neil et al. (eds.), *Teaching and Measuring Cognitive Readiness*,
DOI 10.1007/978-1-4614-7579-8_4, © Springer Science+Business Media New York 2014

does some good for society? So in the process of trying to grow the organization through a new investment, John engages in creative, analytical, practical, and wise thinking.

Second: Qara has been a danger all along. But now intelligence reports make it clear that the country is building nuclear weapons. Qaran officials refuse to allow UN inspections and insist that any nuclear activity in the country is totally for peaceful ends. Moreover, Qara asserts its right to be a nuclear state. It points out that many other countries have gone nuclear, and Qaran officials did not tell these other countries that they should renounce nuclear power. Jane, the president of Amarinth, a peaceful nuclear power, is very concerned about the development of nuclear technology in Qara. For one thing, Qaran officials have openly called for the destruction of Amarinth. For another, Qara borders Amarinth and thus must be considered a threat. And finally, Qara has a history of state-sponsored terrorism, and the government of Amarinth is very worried that Qaran nuclear weapons will be sold or even given to terrorists serving as proxies for Qara. But the Qaran nuclear facilities are scattered so no single strike will entirely disable its nuclear operations. What's to be done?

Jane, working with her advisors, has devised a simultaneous multiple-strike strategy that they all believe could disable 60–80 % of Qara's nuclear facilities in one coordinated blow. The strategy is creative in that it does not rely on any one strike to fulfill the whole goal. Second, Jane and her advisors analyze the strategy carefully, trying their very best to poke holes in it. Third, Jane and advisors ask whether the strategy is practical: Can they actually deliver the strike, is there too much danger of potent retaliation, and what will they do if, or more likely, when world opinion turns against them? Fourth, Jane and her advisors ponder whether the strategy is wise. Will it make the world safer for Amarinth, or actually more dangerous? Will it serve the common good not only of Amarinth, but even of Qara, in the long term? Is it ethical to do such a preemptive strike, or is there some other option? Jane, like John, passes through the creative, analytical, practical, and wisdom-based steps in her thinking process.

Note that all of the steps are important. Without creative thinking, one risks trying again and again old and sometimes inadequate or even failed solutions. Without analytical thinking, one may devise a solution that is novel but that has flaws in it. Without practical thinking, one risks a solution that might work in theory, but not in practice, or a solution that others will not accept. Without wise thinking, one risks a solution that is unethical or that, while working in the short run, does not work in the long run.

So how can we construct a model of instruction and assessment which, from early ages onward, prepares people to think creatively, analytically, practically, and wisely?

Traditional instruction and assessment assume a model of abilities that is narrow and, we will argue, counterproductive. In particular, such models emphasize memory and analytical skills at the expense of other equally important or even more important skills. In this chapter, we propose a broader model that will (a) recognize the skills of a wider band of students, (b) value those skills, and (c) enable individuals as well as educational and societal institutions better to capitalize on these broader skills. The model with which we begin is the augmented theory of successful intelligence.

4.1 Wisdom-Intelligence-Creativity-Synthesized: The Augmented Theory of Successful Intelligence

The augmented theory of successful intelligence suggests that students' failures to achieve at a level that matches their potential often results from teaching and assessment that are narrow in conceptualization and rigid in implementation (Sternberg, 1997, 2003; Sternberg & Grigorenko, 2007; Sternberg, Jarvin, & Grigorenko, 2009). The ways in which educators teach and assess do not always match the ways in which students learn. The traditional methods, in essence, typically shine a spotlight on a small number of students with certain ability-based styles, and almost never focus on a large number of students who have the ability to succeed, but whose ability-based styles do not correspond to the patterns of learning and thinking valued by the schools. To rectify this situation, one must value other ability-based styles and then change teaching and assessment so that these other ability patterns can lead to success in school.

According to the augmented theory of successful intelligence, successful intelligence is (1) the use of an integrated set of abilities needed to attain success in life, however, an individual defines it, within his or her sociocultural context. People are successfully intelligent by virtue of (2) recognizing their strengths and making the most of them while at the same time recognizing their weaknesses and finding ways to correct or compensate for them. Successfully intelligent people (3) adapt to, shape, and select environments through (4) a balanced use of their analytical, creative, practical, and wisdom-based abilities, (5) in an ethical fashion serving a common good (Sternberg, 1997, 1998, 1999, 2003, 2009a). Underlying all four abilities is the role of knowledge, because one cannot think analytically, creatively, practically, or wisely if one does not have knowledge stored in long-term memory to which one can apply one's thinking.

People typically balance the four kinds of abilities. They need creative abilities to generate ideas, analytical abilities to determine whether they are good ideas, practical abilities to implement the ideas and to convince others of the value of those ideas, and wisdom-based abilities to ensure that they are using their other abilities ethically and for a common good. Most people who are successfully intelligent are not equal in these four abilities, but they find ways of making the four abilities work together harmoniously and advantageously.

4.2 Teaching and Assessment Using a Broader Model

All teaching and assessment should be balanced in terms of the ability-based styles they require. At the same time, teachers need to put behind them the false dichotomy between "teaching for thinking" and "teaching for the facts," or between emphases on thinking and emphases on memory. Thinking always requires memory and the knowledge base that is accessed through the use of memory. One cannot analyze what one knows if one knows nothing. One cannot creatively go beyond the existing boundaries of knowledge if one cannot identify those boundaries. One cannot apply what one knows in a practical manner if one does not know anything to apply. And one cannot use one's knowledge ethically if one has no knowledge.

It is for these reasons that we encourage teachers to teach and assess achievement in ways that enable students to analyze, create with, and apply their knowledge, as well as to use their knowledge wisely. When students think to learn, they also learn to think. And there is an added benefit to this method: Students who are taught analytically, creatively, practically, and for wisdom perform better on assessments, apparently without regard to the form the assessments take. That is, they outperform students instructed in conventional ways, even if the assessments are for straight factual memory (Sternberg, Torff, & Grigorenko, 1998a, 1998b).

What, exactly, are the techniques used to teach analytically, creatively, practically, and wisely?

1. *Teaching analytically means encouraging students to* (*a*) *analyze*, (*b*) *critique*, (*c*) *judge*, (*d*) *compare and contrast*, (*e*) *evaluate*, *and* (*f*) *assess*. When teachers refer to teaching for "critical thinking," they typically mean teaching for analytical thinking. How does such teaching translate into instructional and assessment activities? Various examples across the school curriculum are shown in Table 4.1.
2. *Teaching creatively means encouraging students to* (*a*) *create*, (*b*) *invent*, (*c*) *discover*, (*d*) *imagine if…*, (*e*) *suppose that…*, *and* (*f*) *predict*. Teaching for creativity requires teachers not only to support and encourage creativity but also

Table 4.1 Examples of analytical teaching

Example activity	Curriculum area
Analyze the development of the character of Heathcliff in *Wuthering Heights*	Literature
Critique the design of the experiment (just gone over in class or in a reading) showing that certain plants grow better in dim light than in bright sunlight	Biology
Judge the artistic merits of Roy Lichtenstein's "comic-book art," discussing its strengths as well as its weaknesses as fine art	Art
Compare and contrast the respective natures of the American Revolution and the French Revolution, pointing out both ways in which they were similar and ways in which they were different	History
Evaluate the validity of the following solution to a mathematical problem, and discuss its weaknesses, if there are any	Mathematics
Assess the strategy used by the winning player in the tennis match you just observed, stating what techniques she used to defeat her opponent	Physical education

Table 4.2 Examples of creative teaching

Example activity	Curriculum area
Create an alternative ending to the short story you just read that presents a different way things might have gone for the main characters in the story	Literature
Invent a dialogue between an American tourist in Paris and a French man he encounters on the street from whom the American is asking directions on how to get to Rue Pigalle	French
Discover the fundamental physical principle that underlies all of the following problems, each of which differs from the others in the "surface structure" of the problem but not in its "deep structure…"	Physics
Imagine that the government of China keeps evolving over the next 20 years in much the same way it has been evolving. What do you imagine the government of China will be like in 20 years?	Government/ political science
Suppose that you were to design a new instrument to be played in a symphony orchestra for future compositions. What might that instrument be like, and why?	Music
Predict changes that are likely to occur in the vocabulary or grammar of spoken Spanish in the border areas of the Rio Grande over the next 100 years as a result of the continuous interactions between Spanish and English speakers	Linguistics

to role model it and to reward it when it is displayed (Sternberg & Grigorenko, 2007; Sternberg & Lubart, 1995; Sternberg & Spear-Swerling, 1996; Sternberg & Williams, 1996). In other words, teachers not only need to talk the talk, but also to walk the walk. Consider, in Table 4.2, some examples of instructional and assessment activities that encourage students to think creatively.

3. *Teaching practically means encouraging students to* (*a*) *apply*, (*b*) *use*, (*c*) *put into practice*, (*d*) *implement*, (*e*) *employ*, *and* (*f*) *render practical what they know*. Such teaching must relate to the real practical needs of the students, not just to what would be practical for other individuals (Sternberg et al., 2000). Table 4.3 gives examples.
4. *Teaching for wisdom means encouraging students to* (*a*) *apply their knowledge to a common good*, (*b*) *over the long- and short terms*, (*c*) *through the infusion of positive ethical values*, (*d*) *by balancing intrapersonal* (*one's own*), *interpersonal* (*others'*), *and extrapersonal* (*larger*) *interests*. Examples can be found in Table 4.4.

What are some data regarding such teaching and assessment? We first present data on teaching, then on assessment. Because wisdom was added to the theory later, it is not represented in many of our empirical studies.

4.3 Some Data on Instruction

We have sought to test the theory of successful intelligence in the classroom, looking at whether teaching in different ways makes a difference in learning. Consider some studies.

Table 4.3 Examples of practical teaching

Example activity	Curriculum area
Apply the formula for computing compound interest to a problem people are likely to face when planning for retirement	Economics, math
Use your knowledge of German to greet a new acquaintance in Berlin	German
Put into practice what you have learned from teamwork in football to making a classroom team project succeed	Athletics
Implement a business plan you have written in a simulated business environment	Business
Employ the formula relating distance, rate, and time to compute a distance	Math
Render practical a proposed design for a new building that will not work within the aesthetic context of the surrounding buildings, all of which are at least 100 years old	Architecture

Table 4.4 Examples of wisdom-based teaching

Example activity	Curriculum area
Can a war ever *promote a common good*?	Political science
Can the use of stem cells from embryos ever *be based upon positive ethical values*?	Biology
Does a massive economic stimulus package, resulting in a great increase in national debt, *promote short-term interests at the expense of long-term ones*?	Economics
Was the bombing of Hiroshima *ethically justified*?	History
Can one argue that creating a weapon of mass destruction can *serve an extrapersonal good*—that is, a good for the preservation of a society?	Engineering
Should euthanasia be legalized under certain circumstances as *balancing intrapersonal with extrapersonal interests*?	Public health

4.3.1 Aptitude–Treatment Interaction at the High School Level

In a first set of studies, investigators explored the question of whether conventional education in school systematically discriminates against children with creative and practical strengths (Sternberg & Clinkenbeard, 1995; Sternberg, Ferrari, Clinkenbeard, & Grigorenko, 1996; Sternberg, Grigorenko, Ferrari, & Clinkenbeard, 1999). Motivating this work was the belief that systems in most schools strongly favor children with strengths in memory and analytical abilities.

Sternberg, Ferrari, Clinkenbeard, and Grigorenko (1999) used the Sternberg Triarchic Abilities Test (STAT) (Sternberg, 1993), an early measure of successful intelligence, to sort students according to analytical, creative, and practical ability-based styles. The test contained items measuring analytical, creative, and practical abilities, each in the verbal, quantitative, and figural domains. Assessment in the verbal domain was done using both multiple-choice and essay items. The test was administered to 326 children around the United States and in some other countries who were identified by their schools as gifted by any standard whatsoever.

Children were selected for a summer program in (college-level) psychology if they fell into one of the five ability groupings: high analytical, high creative, high practical, high balanced (high in all three abilities), or low balanced (low in all three abilities). These students, who came to Yale, were then divided into four instructional groups. Students in all four instructional groups used the same introductory-psychology textbook (a preliminary version of Sternberg, 1995) and listened to the same psychology lectures. What differed among them was the type of afternoon discussion section to which they were assigned. Each was assigned to an instructional condition that emphasized memory, analytical, creative, or practical instruction. For example, in the memory condition, they might be asked to describe the main tenets of a major theory of depression. In the analytical condition, they might be asked to compare and contrast two theories of depression. In the creative condition, they might be asked to formulate their own theory of depression. In the practical condition, they might be asked how they could use what they had learned about depression to help a friend who was depressed.

Students in all four instructional conditions were evaluated in terms of their performance on homework, a midterm exam, a final exam, and an independent project. Each type of work was evaluated for memory, analytical, creative, and practical quality. Thus, all students were evaluated in exactly the same way.

The results suggested the utility of the theory of successful intelligence. This utility showed itself in several ways.

First, the authors observed that when the students arrived at Yale, those in the high creative and high practical groups were much more diverse in terms of racial, ethnic, socioeconomic, and educational backgrounds than were the students in the high analytical group, suggesting that correlations of measured intelligence with status variables such as these may be reduced by using a broader conception of intelligence. Thus, the kinds of students identified as strong in creative and practical abilities differed in terms of the populations from which they were drawn compared with students identified as strong solely by analytical measures. More importantly, just by expanding the range of abilities measured, the investigators discovered intellectual strengths that might not have been apparent through a conventional test.

Second, Sternberg et al. (1999) found that all three ability tests—analytical, creative, and practical—significantly predicted course performance. When multiple-regression analysis was used, at least two of these ability measures contributed significantly to the prediction of each of the measures of achievement. One of the significant predictors was always the analytical score, perhaps reflecting the difficulty of deemphasizing the analytical way of teaching.

Third and most importantly, there was an aptitude–treatment interaction whereby students who were placed in instructional conditions that better matched their pattern of abilities outperformed students who were mismatched. The assessments thus fit the definition of ability-based styles proposed here. In other words, when students are taught in a way that fits the way they think, they do better in school. Children with creative and practical abilities, who are almost never taught or assessed in a way that matches their pattern of abilities, may be at a disadvantage in course after course, year after year.

4.3.2 Teaching for Successful Intelligence in Science and Social Studies at Grades 3 and 8

If students learn best in different ways, then teaching analytically, creatively, and practically should benefit students overall because it will enable more students to capitalize on strengths and to correct or compensate for weaknesses. A follow-up study (Sternberg et al., 1998a, 1998b) examined learning of social studies and science by third graders and eighth graders. The 225 third graders were students in a very low-income neighborhood in Raleigh, North Carolina. The 142 eighth graders were largely middle to upper-middle class students studying in Baltimore, Maryland, and Fresno, California. In this study, students were assigned to one of the three instructional conditions. In the first condition, they were taught basically the same course they would have received had there been no intervention. The emphasis in the course was on memory. In a second condition, students were taught in a way that emphasized critical (analytical) thinking. In the third condition, they were taught in a way that emphasized analytical, creative, and practical thinking. All students' performance was assessed for memory learning (through multiple-choice assessments) as well as for analytical, creative, and practical learning (through performance assessments).

As expected, students in the successful-intelligence (analytical, creative, practical) condition outperformed the other students in terms of the performance assessments. One could argue that this result merely reflected the way they were taught. Nevertheless, the result suggested that teaching for these kinds of thinking succeeded. More important, however, was the result that children in the successful-intelligence condition outperformed the other children even on the multiple-choice memory tests. In other words, to the extent that one's goal is just to maximize children's memory for information, teaching for successful intelligence is still superior. It enables children to capitalize on their strengths and to correct or to compensate for their weaknesses, and it allows children to encode material in a variety of interesting ways.

4.3.3 Teaching Middle and High School Reading for Successful Intelligence

Grigorenko, Jarvin, and Sternberg (2002) extended these results to reading curricula at the middle school and high school level. In a study of 871 middle school students and 432 high school students, participants were taught language arts either for successful intelligence or through the regular curriculum. At the middle-school level, reading was taught explicitly. At the high school level, reading was infused into instruction in mathematics, physical sciences, social sciences, English, history, foreign languages, and the arts. In all settings, students who were taught for successful intelligence substantially outperformed students taught in standard ways (Grigorenko et al., 2002). Allowing students to bring their styles of learning to bear upon their work increases academic performance.

4.3.4 Teaching Mathematics to Eskimo High School Students

These principles work even in difficult-to-teach populations. In a study in Alaska of Yup'ik Eskimo children, researchers found that children who were taught using practical instruction involving fish-racks, a part of their everyday lives, learned principles of perimeter and area better than did students who were taught using conventional textbook instruction (Sternberg, Lipka, Newman, Wildfeuer, & Grigorenko, 2007). For practical learners, practical teaching allows them to learn in a style that fits them and thus improves academic achievement.

4.3.5 Teaching for Successful Intelligence in Language Arts, Mathematics, and Science in Grade 4

The pattern of results indicating the advantage of teaching for successful intelligence has been replicated in yet another study, the largest so far (Sternberg et al., 2007; see also Sternberg, Grigorenko, & Zhang, 2008; Sternberg, Jarvin, & Grigorenko, 2011). This study was carried out on a national scale and involved thousands of fourth-grade students. As this study is quite recent, I report it in more detail.

In this study, a group of educators and psychologists collaborated to develop and improve instructional materials and assessments in three subject areas for students in fourth grade language arts, mathematics, and science. In addition, this study was also characterized by a conservative experimental design; specifically, curricula based on ideas of teaching for successful intelligence were compared with curricula based on modern theories of memory and critical thinking.

In each of these subject-matter areas, we developed several curriculum units, covering approximately a 12-week classroom intervention period. There were fewer science units than mathematics and language-arts units, but each one of these units covered a longer period. Each unit was composed of a pre-intervention assessment, a teacher guide, a set of activities and materials for students, and a post-intervention assessment.

Each of the units in each subject area was developed in three versions, corresponding to the three educational methods (successful intelligence, critical thinking, and memory) being compared in this study. The three versions were parallel and shared the same knowledge content, but adopted different theoretical foci to teach the content.

The pre- and post-intervention assessments consisted of a set of 30 items (half multiple-choice and half open-ended) related to the unit's content. These assessments were identical for students in all three conditions. In addition to the unit-specific assessments, a general baseline assessment was administered to all students participating in the program (either the Woodcock-Johnson III Test of Achievement, a normed test of academic achievement, or the Ohio Department of Education Proficiency Test for Grade 4).

The instructional materials consisted of a teacher guide containing background content information and instructional guidelines and an activity workbook for students. The activities were labeled according to their level of difficulty (from less challenging to more challenging), and teachers selected those activities they judged best fit the abilities of their students. Particular care was exercised to ensure that the content taught in the three versions of a unit was comparable. The content of the materials was carefully aligned with national and state (for states participating in this study) standards.

Overall, 196 teachers and 7,702 students participated in the study. The study spanned 4 years, 9 states, 14 school districts, and 110 schools. The sample included primarily fourth graders, but also third and fifth graders who were taught by teachers participating in the study with their fourth graders. The number of participants was approximately equal in all experimental groups.

The analyses generally proceeded in two phases. First, multi-facet Rasch analysis was employed to determine scale and item characteristics, inter-rater reliability, and student-ability estimates. Second, student ability estimates were subjected to hierarchical-linear modeling analyses to compare performance for each student in each condition.

Each pre- and post-instructional assessment contained a mix of multiple-choice and open-ended items assessing creative, practical, analytic, and memory abilities specific to each unit. Open-ended items were scored by trained raters using test-specific rubrics. The many-facets Rasch model (FACETS) was applied to derive ability scores that adjusted for differences in rater severity and to equate the difficulty of the pretest and posttest by anchoring items common to each test to the item's calibrated difficulty at posttest.

In addition to the *total test* ability score, subscale scores for the *creative–practical* and *analytical–memory* components were also derived for each participating student. The first score represents ability-based styles proposed in particular by the theory of successful intelligence, whereas the second score represents ability-based styles proposed by conventional theories of intelligence (Sternberg, 1997). This derivation of scores was done in a second series of FACETS analyses that used the calibrated item-difficulty estimates from the first analysis (i.e., on all items) to anchor the difficulty of the subscale items (i.e., separate analyses were run for each subset of items). This means that an individual's pretest and posttest scores on the *creative–practical*, *analytical–memory*, and *total test* are comparable, and more importantly, that gain scores are meaningful. This process was repeated separately for each unit and hence all scores are on the logit scale.

Hierarchical linear modeling was used to compare the gain from pretest to posttest across the three instructional conditions for each unit, separately for each year of implementation of the study. A three-level model was used to predict posttest scores, with the first level corresponding to individual growth from time one (pretest) to time two (posttest), the second level corresponding to students (demographic characteristics and baseline assessment indicators), and the third level corresponding to teachers. Experimental condition was modeled at the teacher level. Estimated gain for a participating student was considered to be the value of the slope when predicting the posttest from the pretest.

Finally, we completed a summative analysis of both samples, using a method initially proposed by Fisher (1954) for combining results from multiple samples to the presence of effect.

The item-response reliability estimates of the *total test* and *creative–practical* scores for each assessment were satisfactory (means = 0.87 and 0.83, medians = 0.87 and 0.83, respectively—coincidentally, the same values to two decimal places). Interestingly, the reliability coefficients for the *analytical–memory* subscales and the standardized baseline measures were consistently lower, but still within acceptable ranges (means = 0.70 and 0.70, medians = 0.74 and 0.70, respectively).

To summarize the patterns of results across the two samples, we applied Fisher's product criterion, which allows unifying different *p*-values obtained for different units within different academic domains across multiple years of study. When considered across multiple units and the three academic domains (language arts, mathematics, and sciences), the *total test* scores of students in the successful-intelligence group were higher than those in the critical-thinking group ($p < 0.01$) and the memory group ($p < 0.05$). When considered separately for *creative–practical* and *analytical–memory* scores, however, the patterns of results were different across experimental conditions. Specifically, children in the successful-intelligence group did better on creative and practical items than did students in the critical-thinking ($p < 0.001$) or memory ($p < 0.001$) groups. Yet, when performance on analytical and memory items was considered, although the students in the successful-intelligence group differed from their peers in the critical-thinking group ($p < 0.001$), they did not differ from participants in the memory group ($p = 0.268$). The abilities thus once again met our definition of ability-based styles.

All three instructional conditions demonstrated substantial gain from pretest to posttest. Yet, the results indicate that students from the successful-intelligence group overall tended to have consistently higher gain scores than did students in the control conditions.

Thus the results of these sets of studies suggest that children learn and think with different ability-based styles. Teaching in ways that recognize these individual differences improves learning.

4.4 Some Data on Ability-Based Styles in High-Stakes Assessments

4.4.1 The Rainbow Project

The Rainbow Project and related collaborations are fully described elsewhere in Sternberg and The Rainbow Project Collaborators (2006; see also Sternberg, 2009b, 2010; Sternberg & The Rainbow Project Collaborators, 2005; Sternberg, The Rainbow Project Collaborators, & The University of Michigan Business School Collaborators, 2004). The Rainbow measures supplement the SAT. The SAT is a 3-h examination currently measuring verbal comprehension and mathematical thinking skills, with a writing component used by some, but not other colleges. A wide variety of studies have shown the utility of the SAT as a predictor of college success, especially as measured by GPA (grade-point average). The Rainbow Project was done before wisdom was added to the model, so it assessed only analytical, creative, and practical skills.

Available data suggest reasonable predictive validity for the SAT in predicting college performance (e.g., Hezlett et al., 2001; Kobrin, Camara, & Milewski, 2002). Indeed, traditional intelligence or aptitude tests have been shown to predict performance across a wide variety of settings. But as is always the case for a single test or type of test, there is room for improvement. The theory of successful intelligence provides one basis for improving prediction and possibly for establishing greater equity and diversity. It suggests that broadening the range of skills tested to go beyond analytical skills to include practical and creative skills as well might significantly enhance the prediction of college performance beyond current levels. Thus, the theory does not suggest *replacing* but rather *augmenting* the SAT in the college admissions process. A collaborative team of investigators sought to study how successful such an augmentation could be.

In the Rainbow Project (Sternberg & The Rainbow Project Collaborators, 2006), data were collected at 15 schools across the United States, including eight 4-year colleges, five community colleges, and two high schools. The participants received either course credit or money. They were 1,013 students predominantly in their first year of college or their final year of high school. In this report, analyses only for college students are discussed because they were the only ones for whom the authors had data available regarding college performance. The final number of participants included in these analyses was 793.

Baseline measures of standardized test scores and high school GPA average were collected to evaluate the predictive validity of current tools used for college admission criteria, and to provide a contrast for the current measures. Students' scores on standardized college entrance exams were obtained from the College Board.

Measuring analytical skills. The measure of analytical skills was provided by the SAT plus analytical items of the STAT (Sternberg, 1993).

Measuring creative skills. Creative skills were measured by STAT multiple-choice items and by performance-based items. Creative skills also were measured using open-ended measures. One measure required writing two short stories with a selection from among unusual titles, such as "The Octopus's Sneakers," one required orally telling two stories based on choices of picture collages, and the third required captioning cartoons from among various options. Open-ended performance-based answers were rated by trained raters for novelty, quality, and task-appropriateness. Multiple judges were used for each task and satisfactory reliability was achieved (details in Sternberg & The Rainbow Project Collaborators, 2006).

Measuring practical skills. Multiple-choice measures of practical skills were obtained from the STAT. Practical skills also were assessed using three situational-judgment inventories: the Everyday Situational Judgment Inventory (Movies), the Common Sense Questionnaire, and the College Life Questionnaire, each of which taps different types of tacit knowledge. The general format of tacit-knowledge inventories has been described in Sternberg et al. (2000), so only the content of the inventories used in this study is described here. The movies present everyday situations that confront college students, such as asking for a letter of recommendation from a professor who shows through nonverbal cues that he does not recognize the student very well. One then has to rate various options for how well they would work in response to each situation. The Common Sense Questionnaire provides everyday business problems, such as being assigned to work with a coworker whom one cannot stand. The College Life Questionnaire provides everyday college situations for which a solution is required.

Unlike the creativity performance tasks, in the practical performance tasks the participants were not given a choice of situations to rate. For each task, participants were told that there was no “right” answer, and that the options described in each situation represented variations on how different people approach different situations.

An example of a creative item might be to write a story using the title “3516” or “It’s Moving Backward.” Another example might show a collage of pictures in which people are engaged in a different wide variety of activities helping other people. One would then orally tell a story that takes off from the collage. An example of a practical item might show a movie in which a student has just received a poor grade on a test. His roommate had a health crisis the night before, and he had been up all night helping him. His professor hands him back the test paper, with a disappointed look on her face, and suggests to the student that he study harder next time. The movie then stops. The student then has to describe how he would handle the situation. Or the student might receive a written problem describing a conflict with another individual with whom he is working on a group project. The project is getting mired down in the interpersonal conflict. The student has to indicate how he would resolve the situation to get the project done.

Administrative details. All materials were administered in one of the two formats. A total of 325 of the college students took the test in paper-and-pencil format; 468 students took the test on the computer via the World Wide Web. Participants were tested either individually or in small groups. During the oral stories section, participants who were tested in groups either wore headphones or were directed into a separate room so as not to disturb the other participants during the story dictation.

Basic data. When examining college students alone, this sample showed a slightly higher mean level of SAT than that found in colleges across the country. The sample means on the SATs were, for 2-year college students, 490 verbal and 508 math, and for 4-year college students, 555 verbal and 575 math. These means, although slightly higher than typical, were within the range of average college students.

There is always a potential concern about restriction of range in scores using the SAT when considering students from a select sample of universities, especially when the means run a bit high. Restriction of range means that one tests a narrower range of student skill levels than that which is representative of the entire population that actually takes the SAT. However, the sample was taken from a wide range in selectivity of institutions, from community colleges to highly select 4-year institutions. In fact, statistics assessing range showed that the sample ranged somewhat *more* widely than is typical for the test. Because there was no restriction of range, there was no need to correct for it.

Another potential concern is pooling data from different institutions. We pooled data because in some institutions we simply did not have large enough numbers of cases for the data to be meaningful.

Factor structure of the Rainbow measures. Some scholars believe that there is only one set of skills that is highly relevant to school performance, what is sometimes called "general ability," or *g* (e.g., Jensen, 1998). These scholars believe that tests may appear to measure different skills, but when statistically analyzed, show themselves merely to measure the single general ability. Does the test actually measure distinct analytical, creative, and practical skill groupings? Factor analysis addresses this question. Three meaningful factors were extracted from the data. One factor represented practical performance tests. The second, a weaker factor, represented the creative performance tests. The third factor represented the multiple-choice tests (including analytical, creative, and practical). Thus, method variance proved to be very important. The results show the importance of measuring ability-based styles using multiple formats, precisely because method is so important in determining factorial structure.

Predicting College GPA. College admission officers are not interested, exactly, in whether these tests predict college success. Rather, they are interested in the extent to which these tests predict college success *beyond* those measures currently in use, such as the SAT and high school GPA. To test the incremental validity provided by Rainbow measures above and beyond the SAT in predicting GPA, a series of statistical analyses (called hierarchical regressions) was conducted that included the items analyzed in the analytical, creative, and practical assessments.

If one looks at the simple correlations, the SAT-V, SAT-M, high school GPA, and the Rainbow measures all predict freshman-year GPA. But how do the Rainbow measures fare on incremental validity? In one set of analyses, the SAT-V, SAT-M, and high school GPA were included in the first step of the prediction equation because these are the standard measures used today to predict college performance. Only high school GPA contributed uniquely to prediction of college GPA. In Step 2, the analytic subtest of the STAT was added, because this test is closest conceptually to the SAT tests. The analytical subtest of the STAT slightly but significantly increased the level of prediction. In Step 3, the measures of practical ability were added, resulting in a small increase in prediction. The inclusion of the creative measures in the final step of this prediction equation indicates that, by supplementing the SAT and high school GPA with measures of analytical, practical, and creative abilities, a total of 24.8 % of the variance in GPA can be accounted for. Inclusion of the Rainbow measures in Steps 2, 3, and 4 represents an increase of about 9.2 % (from 0.156 to 0.248) in the variance accounted for over and above the typical predictors of college GPA. Including the Rainbow measures without high school GPA, using only SAT scores as a base, represents an increase in percentage variance accounted for of about 10.1 % (from 0.098 to 0.199). Looked at in another way, this means that the Rainbow measures roughly doubled prediction vs. the SAT alone. Different ability-based styles of thinking, then, make a difference in predicting academic achievement beyond unitary measures of traditional general ability.

These results suggest that the Rainbow tests add considerably to the prediction gotten by SATs alone. They also suggest the power of high school GPA, particularly, in prediction, because it is an atheoretical composite that includes within it many variables, including motivation and conscientiousness.

Group differences. Although one important goal of the present study was to predict success in college, another important goal involved developing measures that reduce ethnic group differences in mean levels. There has been a lively debate as to why there are socially defined racial group differences, and as to whether scores for members of underrepresented minority groups are over- or under-predicted by SATs and related tests (see, e.g., Bowen & Bok, 2000). Might it be because different ethnic groups, on average, show different ability-based styles of thinking as a result of differential socialization? There are a number of ways one can test for group differences in these measures, each of which involves a test of the size of the effect of ethnic group. Two different measures were chosen.

First, consider numbers showing the impact of ethnic group on test scores (called omega squared coefficients). This procedure involves considering differences in mean performance levels among the six ethnic and racial groups reported, including European American, Asian American, Pacific Islander, Latino American, African American, and Native American (American Indian), for the following measures: the baseline measures (SAT-V and SAT-M), the STAT ability scales, the creativity performance tasks, and the practical-ability performance tasks. The coefficient indicates the proportion of variance in the variables that is accounted for by the self-reported ethnicity of the participant. The omega squared values were 0.09 for SAT-V, 0.04 for SAT-M, and 0.07 for combined SAT. For the Rainbow measures, omega squared values ranged from 0.00 to 0.03 with a median of 0.02. Thus, the Rainbow measures showed reduced values relative to the SAT.

Another test of effect sizes (Cohen's *D*) allows one to consider more specifically a representation of specific group differences. For the test of ethnic group differences, each entry represents how far away from the mean for European Americans each group performs in terms of standardized units of test scores (standard deviations). For the test of gender differences, the entries represent how far away women perform from men in terms of standard deviations.

These results indicate two general findings. First, in terms of overall differences, the Rainbow tests appear to reduce ethnic group differences relative to traditional assessments of abilities like the SAT. Second, in terms of specific differences, it appears that the Latino American students benefit the most from the reduction of group differences. African American students, too, seem to show a reduction in difference from the European American mean for most of the Rainbow tests, although a substantial difference appears to be maintained with the practical performance measures. Important reductions in differences can also be seen for the Native

American students relative to European American students. Indeed, their median was higher for the creative tests. However, the very small sample size suggests that any conclusions about Native American performance should be made tentatively.

Although the group differences are not perfectly reduced, these findings suggest that measures can be designed that reduce ethnic and socially defined racial group differences on standardized tests, particularly for historically disadvantaged groups like African American and Latino American students. These findings have important implications for reducing adverse impact in college admissions. They suggest that different groups do have, on average, different patterns of ability-based styles. Similar findings have been obtained by Fagan and Holland (2007), who found that differences in scores on ability tests were due in large part to race-specific knowledge.

4.4.2 Data from Other Assessment Projects

The principles behind the Rainbow Project apply at other levels of admissions as well. For example, Hedlund, Wilt, Nebel, Ashford, and Sternberg (2006) have shown that the same principles can be applied in admissions to business schools, also with the result of increasing prediction and decreasing ethnic- (as well as gender-) group differences by including tests of practical thinking in addition to the Graduate Management Admission Test (GMAT).

Stemler, Grigorenko, Jarvin, and Sternberg (2006) studied measurement of ability-based styles in the context of achievement testing (see also Stemler, Sternberg, Grigorenko, Jarvin, & Sharpes, 2009). In this project, funded by the Educational Testing Service and the College Board, they asked whether the same principles could be applied to high-stakes achievement testing used for college admissions and placement. They modified Advanced Placement tests in psychology and statistics additionally to assess analytical, creative, and practical skills. Here is an example in psychology.

A variety of explanations have been proposed to account for why people sleep.

(a) Describe the restorative theory of sleep (*memory*).
(b) An alternative theory is an evolutionary theory of sleep, sometimes referred to as the “preservation and protection” theory. Describe this theory and compare and contrast it with the restorative theory. State what you see as the two strong points and two weak points of this theory compared to the restorative theory (*analytical*).
(c) How might you design an experiment to test the restorative theory of sleep? Briefly describe the experiment, including the participants, materials, procedures, and designs (*creative*).
(d) A friend informs you that she is having trouble sleeping. Based on your knowledge of sleep, what kinds of helpful (and health-promoting) suggestions might you give her to help her fall asleep at night (*practical*)?

As in the other studies, the investigators found that by asking such questions, they were able both to increase the range of skills they tested and substantially to reduce ethnic group differences in test scores. Again, different ethnic groups seem to show different modal patterns of ability-based styles.

In collaboration with our colleagues from a private preparatory school (Grigorenko et al., 2009), we developed a supplementary battery of admission assessments, which, in addition to taking into account students' SSAT (Secondary School Admission Test) scores, allowed this school to consider students' creative and practical styles. Specifically, we developed two assessments of practical competence (style) and two assessments of creative competence (style). One practical-competence task surveyed students' readiness to adapt to the new environment of a boarding school and navigate the new "rules and regulations" of a highly academically oriented and demanding prep school. In this task, students were expected to rate a number of solutions offered to them after they read a description of a practical situation. The second task included more generic situations descriptive of social aspects of student life. In this assessment, a problematic situation was depicted and participants were asked to continue the story by identifying with the main character and developing the next step in the plot. Creative competence was also assessed with two different tasks. One task asked for a brief story under one of the five proposed titles: (1) Too Much, Too Fast; (2) The Landing on the Planet Vespa; (3) Third Time's the Charm; (4) The Spy Was Not Captured After All; and (5) When the Music Stopped. The second task included different word problems describing various situations related to novel uses of scientific knowledge; students were asked to find a solution using some knowledge of the sciences. These four indicators were used in regression analyses predicting freshman GPA for a class of 152 students. When introduced into regression after SSAT Verbal, Quantitative, and Reading indicators, the practical-competence tasks doubled the prediction (from 12.0 to 24.4 %) and the creative-competence tasks added an additional 4.4 % (from 24.4 to 28.8 %).

Thus, tests such as the Rainbow Assessment do not benefit only members of ethnic minority groups. There are many students who come from the majority group, and even from well-off homes, who learn in ways that are different from those assessed by conventional standardized tests. These children may well have the abilities they need to succeed in life and even in school; but these abilities may not be reflected in scores on conventional tests. Our tests help identify such students.

It is one thing to have a successful research project, and another actually to implement the procedures in a high-stakes situation. Can any of these ideas actually make a difference in practice?

4.5 Practical Implementation

Tufts University strongly emphasizes the role of active citizenship in education. So it seemed like a suitable setting to put into practice some of the ideas from the Rainbow Project. Tufts instituted the Kaleidoscope Project, which represents an implementation of the ideas of Rainbow, but goes beyond that project to include in its assessment the construct of wisdom (Sternberg, 2007a, 2007b).

Tufts placed questions designed to assess wisdom, intelligence, and creativity synthesized (Wisdom-Intelligence-Creativity-Synthesized [WICS]—Sternberg, 2003, 2007a) on the 2006–2007 application for all of the more than 15,000 students applying for undergraduate admissions to Arts, Sciences, and Engineering at Tufts. The questions were optional. Whereas the Rainbow Project was done as a separate high-stakes test administered with a proctor, the Kaleidoscope Project (Sternberg, 2009b; Sternberg, Bonney, Gabora, & Merrifield, 2012; Sternberg et al., 2010) was done as a section of the Tufts-specific part of the college application. It just was not practical to administer a separate high-stakes test such as the Rainbow assessment for admission to one university. Moreover, the advantage of Kaleidoscope is that it moved Tufts away from the high-stakes testing situation in which students must answer complex questions in very short amounts of time under incredible pressure. The section was optional this past year, and students were encouraged to answer just a single question. As examples, a creative question asked students to write stories with titles such as "The End of MTV" or "Confessions of a Middle-School Bully." Another creative question asked students what the world would be like if some historical event had come out differently, for example, if Rosa Parks had given up her seat on the bus. Yet another creative question, a nonverbal one, gave students an opportunity to design a new product or an advertisement for a new product. A practical question queried how students had persuaded friends of an unpopular idea they held. A wisdom question asked students how a passion they had could be applied toward a common good.

So what happened? Some stakeholders were afraid that numbers of applications would go down; instead, they went up. Notably, the quality of applicants rose substantially. There were notably fewer students in what had previously been the bottom third of the pool in terms of quality. Many of those students, seeing the new application, decided not to bother to apply. Many more strong applicants applied. Other stakeholders were concerned that average SATs would go down and perhaps even plummet. Instead, they went up, rising to more than 1,400 (V + M) for the first time. The reason is that the new assessments are not negatively correlated with SATs. Rather, they just are not much correlated at all, one way or another. The squared correlations of the Kaleidoscope assessments with SATs were all less than 0.1. In contrast, squared correlations with quality of extracurricular activities were in the 0.4 range. Merely doing the Kaleidoscope essays had a trivial effect on admission. But students who had an "A" (top rating) on the Kaleidoscope assessments were twice as likely to be admitted as those who did not. The assessments provided a quantified way of assessing ability-based styles of thinking that, in the past, had been assessed only in a more qualitative way. We note that all of these results are correlational, not causal, so that one cannot conclude that Kaleidoscope was the cause of any differences obtained in the past year.

In later analyses, we followed up on students admitted through the Kaleidoscope program in order to compare higher and lower scoring Kaleidoscope students. The higher scoring students excelled in participation in extracurricular and leadership activities. Controlling for high school GPA and SATs, students who were rated for Kaleidoscope outperformed students who were not in freshman GPA. These results were notable given that there were no ethnic group differences in scores on Kaleidoscope!

In sum, adopting these new methods results in the admission of applicants who are more qualified, but in a broader way than was considered in the past. Perhaps most rewarding were the positive comments from large numbers of applicants who completed the essays, irrespective of whether they were later accepted or not, that they felt our application gave them a chance to show themselves for who they are.

After a number of years in which applications by underrepresented minorities were relatively flat in terms of numbers, in 2006–2007 they went up substantially. In the end, Tufts admitted roughly 30 % more African American students than the year before, and 15 % more Latino Americans. So these results, like those of the Rainbow Project, showed that it is possible to increase academic quality and diversity simultaneously, and to do so in for an entire undergraduate class at a major university, not just for small samples of students at some scattered colleges. Most importantly, the university sent a message to students, parents, high school guidance counselors, and others that it believes that there is more to a person than the narrow spectrum of skills assessed by standardized tests, and that these broader skills can be assessed in a quantifiable way.

One might wonder how one assesses answers to questions that seem so subjective. The answer is through well-developed rubrics. For example, we assess analytical responses on the basis of the extent to which they are (a) analytically sound, (b) balanced, (c) logical, and (d) organized. We assess creative responses for how (a) original and (b) compelling they are, as well as on the basis of their (c) appropriateness to the task with which the students were presented. We assess practical responses on the basis of how feasible they are with respect to (a) time, (b) place, and (c) human and (d) material resources. We assess wisdom-based responses on the extent to which they (a) promote a common good by (b) balancing one's own with others' and larger interests, (c) over the long- and short terms, through (d) the infusion of positive (prosocial) values.

4.6 Conclusion

The augmented theory of intelligence provides an integrated model for instruction and assessment that broadens the way we think about abilities. It provides a way for larger numbers of people to succeed, and for schools and society to capitalize on the strengths of all rather than just the few. The model has shown success with both children and adults. For example, Sternberg et al. (2000) showed how it is possible to use situational-judgment assessments to measure practical intelligence in everyday life. But practical intelligence is not enough. The recession of 2008 was created in part by individuals schooled in some of the top universities in the country and then who had a great deal of practical experience on the job. Without teaching for wisdom, societies end up in the kinds of messes that they find themselves in today, but that we hope will be a thing of the past in the future.

Cognitive readiness, on this model, involves more than high IQ or ASVAB score or even high grades in military training. It requires creativity to generate adaptive responses to novel situations, analytical intelligence to ascertain whether the response is indeed a good one, practical intelligence to implement the response and convince others of its value, and wisdom to ensure that the response is toward the common good. In this chapter, I have described various assessments we have created that assess these aspects of cognitive readiness. If we are to have truly cognitively ready personnel, we have to go beyond traditional assessments toward ones that measure a broader range of critical skills.

Acknowledgments The work reported herein was partially supported by a grant from the Office of Naval Research, Award Number N000140810126. The findings and opinions expressed here do not necessarily reflect the positions or policies of the Office of Naval Research.

References

Bowen, W. G., & Bok, D. (2000). *The shape of the river*. Princeton, NJ: Princeton University Press.

Fagan, J. F., & Holland, C. R. (2007). Racial equality in intelligence: Predictions from a theory of intelligence as processing. *Intelligence, 35*, 319–334.

Fisher, R. A. (1954). *Statistical methods for research workers*. New York: Hafner.

Grigorenko, E. L., Jarvin, L., Diffley, R., Goodyear, J., Shanahan, E. J., & Sternberg, R. J. (2009). Are SSATs and GPA enough? A theory-based approach to predicting academic success in high school. *Journal of Educational Psychology, 101*, 964–981.

Grigorenko, E. L., Jarvin, L., & Sternberg, R. J. (2002). School-based tests of the triarchic theory of intelligence: Three settings, three samples, three syllabi. *Contemporary Educational Psychology, 27*, 167–208.

Hedlund, J., Wilt, J. M., Nebel, K. R., Ashford, S. J., & Sternberg, R. J. (2006). Assessing practical intelligence in business school admissions: A supplement to the graduate management admissions test. *Learning and Individual Differences, 16*, 101–127.

Hezlett, S., Kuncel, N., Vey, A., Ones, D., Campbell, J., & Camara, W. J. (2001). *The effectiveness of the SAT in predicting success early and late in college: A comprehensive meta-analysis*. Paper presented at the annual meeting of the National Council of Measurement in Education, Seattle, WA.

Jensen, A. R. (1998). *The g factor*. Westport, CT: Praeger/Greenwood.

Kobrin, J. L., Camara, W. J., & Milewski, G. B. (2002). *The utility of the SAT I and SAT II for admissions decisions in California and the nation* (College Board Report No. 2002-6). New York: College Entrance Examination Board.

Stemler, S. E., Grigorenko, E. L., Jarvin, L., & Sternberg, R. J. (2006). Using the theory of successful intelligence as a basis for augmenting AP exams in psychology and statistics. *Contemporary Educational Psychology, 31*(2), 344–376.

Stemler, S., Sternberg, R. J., Grigorenko, E. L., Jarvin, L., & Sharpes, D. K. (2009). Using the theory of successful intelligence as a framework for developing assessments in AP Physics. *Contemporary Educational Psychology, 34*, 195–209.

Sternberg, R. J. (1993). *Sternberg Triarchic Abilities Test*. Unpublished test. New Haven, CT: Yale University Press.

Sternberg, R. J. (1995). *In search of the human mind*. Ft. Worth, TX: Harcourt.

Sternberg, R. J. (1997). *Successful intelligence*. New York: Plume.

Sternberg, R. J. (1998). Ability testing, instruction and assessment of achievement: Breaking out of the vicious circle. *NASSP Bulletin, 82*, 4–10.

Sternberg, R. J. (1999). The theory of successful intelligence. *Review of General Psychology, 3*, 292–316.

Sternberg, R. J. (2003). *Wisdom, intelligence, and creativity, synthesized*. New York: Cambridge University Press.

Sternberg, R. J., Ferrari, M., Clinkenbeard, P. R., & Grigorenko, E. L. (1996). Identification, instruction, and assessment of gifted children: A construct validation of a triarchic model. *Gifted Child Quarterly, 40*(3), 129–137.

Sternberg, R. J., Grigorenko, E. L., Ferrari, M., & Clinkenbeard, P. (1999). A triarchic analysis of an aptitude–treatment interaction. *European Journal of Psychological Assessment, 15*(1), 1–11.

Sternberg, R. J., & the Rainbow Project Collaborators. (2005). Augmenting the SAT through assessments of analytical, practical, and creative skills. In W. Camara & E. Kimmel (Eds.). *Choosing students: Higher education admission tools for the 21st century* (pp. 159–176). Mahwah, NJ: Lawrence Erlbaum Associates.

Sternberg, R. J. (2007a). Finding students who are wise, practical, and creative. *The Chronicle of Higher Education, 53*(44), B11.

Sternberg, R. J. (2007b). How higher education can produce the next generation of positive leaders. In M. E. Devlin (Ed.), *Futures forum 2007* (pp. 33–36). Cambridge, MA: Forum for the Future of Higher Education.

Sternberg, R. J. (2009a). A new model for teaching ethical behavior. *The Chronicle of Higher Education, 55*(33), B14–B15.

Sternberg, R. J. (2009b). The Rainbow and Kaleidoscope projects: A new psychological approach to undergraduate admissions. *European Psychologist, 14*, 279–287.

Sternberg, R. J. (2010). *College admissions for the 21st century*. Cambridge, MA: Harvard University Press.

Sternberg, R. J., Bonney, C. R., Gabora, L., & Merrifield, M. (2012). WICS: A model for college and university admissions. *Educational Psychologist, 47*(1), 30–41.

Sternberg, R. J., & Clinkenbeard, P. R. (1995). The triarchic model applied to identifying, teaching, and assessing gifted children. *Roeper Review, 17*(4), 255–260.

Sternberg, R. J., Coffin, L., Bonney, C. R., Gabora, L., Jarvin, L., & Karelitz, T. (2010). Broadening the spectrum of undergraduate admissions. *College and University, 86*(1), 2–17.

Sternberg, R. J., Forsythe, G. B., Hedlund, J., Horvath, J., Snook, S., Williams, W. M., et al. (2000). *Practical intelligence in everyday life*. New York: Cambridge University Press.

Sternberg, R. J., & Grigorenko, E. L. (2007). *Teaching for successful intelligence* (2nd ed.). Arlington Heights, IL: Skylight.

Sternberg, R. J., Grigorenko, E. L., Ferrari, M., & Clinkenbeard, P. (1999). A triarchic analysis of an aptitude–treatment interaction. *European Journal of Psychological Assessment, 15*(1), 1–11.

Sternberg, R. J., Grigorenko, E. L., & Zhang, L.-F. (2008). Styles of learning and thinking matter in instruction and assessment. *Perspectives on Psychological Science, 3*(6), 486–506.

Sternberg, R. J., Jarvin, L., & Grigorenko, E. L. (2009). *Teaching for wisdom, intelligence, creativity, and success*. Thousand Oaks, CA: Sage.

Sternberg, R. J., Jarvin, L., & Grigorenko, E. L. (2011). *Explorations of giftedness*. New York: Cambridge University Press.

Sternberg, R. J., Lipka, J., Newman, T., Wildfeuer, S., & Grigorenko, E. L. (2007). Triarchically based instruction and assessment of sixth-grade mathematics in a Yup'ik cultural setting in Alaska. *International Journal of Giftedness and Creativity, 21*(2), 6–19.

Sternberg, R. J., & Lubart, T. I. (1995). *Defying the crowd: Cultivating creativity in a culture of conformity*. New York: Free Press.

Sternberg, R. J., & Spear-Swerling, L. (1996). *Teaching for thinking*. Washington, DC: American Psychological Association.

Sternberg, R. J., & The Rainbow Project Collaborators. (2006). The Rainbow Project: Enhancing the SAT through assessments of analytical, practical and creative skills. *Intelligence, 34*(4), 321–350.

Sternberg, R. J., The Rainbow Project Collaborators, & University of Michigan Business School Project Collaborators. (2004). Theory based university admissions testing for a new millennium. *Educational Psychologist, 39*(3), 185–198.

Sternberg, R. J., Torff, B., & Grigorenko, E. L. (1998a). Teaching for successful intelligence raises school achievement. *Phi Delta Kappan, 79*(9), 667–669.

Sternberg, R. J., Torff, B., & Grigorenko, E. L. (1998b). Teaching triarchically improves school achievement. *Journal of Educational Psychology, 90*, 1–11.

Sternberg, R. J., & Williams, W. M. (1996). *How to develop student creativity*. Alexandria, VA: Association for Supervision and Curriculum Development.

Chapter 5
Situational Load and Personal Attributes: Implications for Cognitive Readiness, Adaptive Readiness, and Training

Stephen J. Zaccaro, Eric J. Weis, Tiffani R. Chen, and Michael D. Matthews

Operational effectiveness in complex domains depends upon the degree of *readiness* individuals bring to performance (Morrison & Fletcher, 2002). While effectiveness refers to an evaluation of actual performance, readiness reflects the "potential of units or individuals to perform well" (Morrison & Fletcher, 2002, p. I-1). Such readiness can refer to the potential of multiple aspects of the person, as well as the conditions of the operational context to actualize that potential. For example, school readiness, or readiness for kindergarten through collegiate educational experiences (Wesley & Buysse, 2003), reflects the state of a student's cognitive, behavioral, social, and motivational preparation for learning and educational performance (Bierman, Torres, Domitrovich, Welsh, & Gest, 2009; Le, Casillas, Robbins, & Langley, 2005; Peterson, Casillas, & Robbins, 2006; Robbins et al., 2004). Workforce readiness refers to the skills workers can bring to effective job performance (O'Neil, Allred, & Baker, 1992). In team and organizational research,

S.J. Zaccaro (✉) • T.R. Chen
Department of Psychology, George Mason University, 4400 University Drive, Fairfax, VA 22030-4444, USA
e-mail: szaccaro@gmu.edu; tiffanirchen@gmail.com

E.J. Weis
Simon Center for the Professional Military Ethic, United States Military Academy, West Point, NY 10996, USA
e-mail: eric.j.weis@us.army.mil

M.D. Matthews
Department of Behavioral Sciences and Leadership, U.S. Military Academy, West Point, NY 10996, USA
e-mail: Mike.Matthews@usma.edu

H.F. O'Neil et al. (eds.), *Teaching and Measuring Cognitive Readiness*,
DOI 10.1007/978-1-4614-7579-8_5,

readiness has been used to reflect a collective's preparedness to adapt to change (Armenakis, Harris, & Mossholder, 1993; Eby, Adams, Russell, & Gaby, 2000). Combat readiness indicates at any chosen point in time the state of individuals in terms of their performance and skill qualifications as well as the presence and state of equipment and other performance resources going into combat missions (Morrison & Fletcher, 2002). In each of these domains, readiness is defined as a critical precursor to effective performance. Moreover, the components of readiness drive the content and need for education, training, and development.

In military domains, research on readiness has focused primarily on the degree of cognitive potential personnel bring to combat missions (Cosenzo, Fatkin, & Patton, 2007; Morrison & Fletcher, 2002; Smyth, 2007). Such *cognitive readiness* was defined by Morrison and Fletcher (2002, p. I-3) as, "the mental preparation (including skills, knowledge, abilities, motivations, and personal dispositions) an individual needs to establish and sustain competent performance in the complex and unpredictable environment of modern military operations." They included the following ability or skill elements as components of cognitive readiness: situation awareness, memory, transfer of training, metacognition, automaticity, problem solving, decision making, mental flexibility, and creativity, leadership, and emotion. These components reflex a mix of cognitive skills (e.g., situation awareness), cognitive processes (e.g., memory), and noncognitive elements (e.g., emotion). Researchers at the Center for Research on Evaluation, Standards, and Student Testing (CRESST; O'Neil, Lang, Perez, Escalante, and Fox, this volume) revised the Morrison and Fletcher's (2002) factors of cognitive readiness more tightly around cognitive-based knowledge, skills, and abilities (KSAs). Thus, they argue that cognitive readiness reflects skills and competencies in adaptability, adaptive expertise, creative thinking, decision making, adaptive problem solving, metacognition, situation awareness, and teamwork. Specifying the components of cognitive readiness more in terms of skills and competencies increases their utility to define and drive training content.

In this chapter we focus on the individual's *readiness to adapt to changing operational and environmental contingencies*. Adaptation was defined both by Morrison and Fletcher (2002) and O'Neil et al. (this volume) as central to the meaning of cognitive readiness. For example, Morrison and Fletcher (2002, p. I-3) noted that "the concept of cognitive readiness may be of special relevance and significance to those who must adapt quickly to emerging, unforeseen challenges." O'Neil et al. (this volume) also included both adaptability and adaptive expertise in their set of cognitive readiness competencies.

We suggest that cognitive readiness can reflect preparation for two distinct levels of performance. The first level refers to the "routine" cognitive aspects of problem situations. Thus, such readiness may reflect an individual's preparation to engage in situation assessment, analysis, and problem solving in familiar or "typical" missions. The second level of cognitive readiness rests on this first level but also reflects preparation to engage in the additional cognitive processes more peculiar to adaptive performance contexts. This distinction mirrors the differences between "routine" and "adaptive" expertise described by several researchers (e.g., Holyoak, 1991; Kozlowski, 1998; Smith, Ford, & Kozlowski, 1997). Routine expertise refers to skill in recognizing and applying well-known rules, procedures, and solutions to typical problems; adaptive expertise reflects skills in understanding when and why existing procedures no longer apply to changing problems and knowing how to adjust problem-solving strategies (Kozlowski, 1998). Different componential cognitive skills likely derive from and contribute to each form of expertise. Likewise, we argue that readiness for complex problem situations that require adaptation centers on competencies that are broader in scope and different in kind than those for more "routine," albeit also complex problems. Accordingly, in this chapter we refer to the former as "adaptive readiness." We will describe the cognitive processes and skills that correspond more closely to adaptive readiness.

Another theme in this chapter argues that adaptive readiness includes more than cognitive preparedness. Many adaptive situations present to individuals not only cognitive demands but also significant emotional and social ones. Indeed, some situations may call for fewer cognitive resources and greater emotional and social capacities. For example, many military combat situations may not necessarily require altering problem strategies, but would require adapting to significant emotional stress (e.g., the wounding or death of fellow soldiers) or adjusting to unfamiliar social situations (e.g., establishing working relationships with tribal or local representatives in foreign countries) (cf., Pulakos, Arad, Donovan, & Plamondon, 2000). Although some level of cognitive demands still exists in such situations, operational effectiveness may rest more strongly on an individual's readiness to deploy emotional or social resources.

While this theme has not been strongly represented in the military readiness literature, similar elements have been defined as part of school readiness. For example, Robbins et al. (2004) included social involvement as a school readiness construct. Le et al. (2005) included such skills in working collaboratively with others and in developing and maintaining relationships with others as part of their Student Readiness Inventory. We also note that Morrison and Fletcher (2002) reported an emotion component to cognitive readiness. However, their focus was on maintaining effective cognitive performance under emotional conditions. Adaptive readiness also may entail skill in adapting more directing to emotional challenges by, for example, maintaining motivational focus and deploying effective coping responses.

In the next section of this chapter, we examine more closely the nature of adaptation, noting in particular the cognitive skills and processes that denote effective adaptive performance. We also describe how adaptive performance demands can present different levels of cognitive, social, and emotional demands on performers. We refer to these demands, respectively, as the situation's cognitive, social, and emotional load. These different demands carry implications for (a) the competencies and skills that define adaptive readiness and (b) the range of training strategies that contribute to growth in adaptive readiness. Several researchers have noted that training strategies needed to grow adaptability skills are different from those used in more traditional training domains (Bell & Kozlowski, 2002, 2008; Ely, Zaccaro, & Conjar, 2009; Kozlowski et al., 2001; Nelson, Zaccaro, & Herman, 2010; Smith et al., 1997). However, we would argue that most adaptability training strategies focus on the cognitive skills necessary for adaptation, i.e., those that promote cognitive readiness for adaptive performance. Later in this chapter, we augment these strategies with ones that promote emotional and sozial readiness for adaptive performance as well.

5.1 The Nature of Adaptive Performance and the Components of Adaptive Readiness

Adaptive performance has been defined as distinct from other forms such as contextual or task performance (Pulakos et al., 2000). Most definitions of adaptation, whether at the individual or team level, describe the core of such performance as reflecting a change in behavior or performance strategies to realign with changed conditions in the operational environment (Banks, Bader, Fleming, Zaccaro, & Barber, 2001; Chan, 2000; Ely et al., 2009; LePine, Colquitt, & Erez, 2000; White et al., 2005). At the team level, adaptation refers to members adjusting their collaborative processes, role-based relationships, or collective performance strategies as the team's environment changes (Burke, Stagl, Salas, Pierce, & Kendall, 2006; Kozlowski, Gully, Nason, & Smith, 1999; LePine, 2003). Ely et al. (2009) emphasized two aspects of the definitions in this literature that are particularly relevant for adaptive readiness. First, adaptation is a *functional* response to environment change, meaning that strategy adjustments foster improved performance—"changes in behavior that are not effective under the new environmental conditions are not considered adaptive" (Ely et al., 2009, p. 176; see also Banks et al., 2001; White et al., 2005). Because notions of readiness reflect one's potential for effective performance, this functionality means that adaptive readiness includes the capacity to identify and enact the performance strategies most likely to bring the individual or unit into realignment with altered environments. Herein lies one of the conditions for a cognitive aspect of adaptive readiness. Cognitive readiness with respect to adaptation includes competencies in matching problem solutions and strategies to appropriate environmental circumstances. This element of adaptive readiness is similar to adaptive expertise, defined as knowing how and when particular solutions will or will not work across different problem domains (Kozlowski, 1998; Smith et al., 1997; see also Ericsson, this volume).

A second point of distinction highlighted by Ely et al. (2009) argues that adaptation does not merely represent shifts in one's level or amount of current responding; instead it reflects a qualitative shift to an entirely different performance strategy (Chan, 2000). This point about adaptation reflects both cognitive and behavioral components of adaptive readiness. Adaptation may sometimes require the development of innovative or novel performance strategies, when those in one's behavioral repertoire are insufficient for new environmental challenges (Burke et al., 2006). Thus, adaptive performance may reflect the application of creative-thinking and problem-solving skills, a component of cognitive readiness defined by both Morrison and Fletcher (2002) and O'Neil et al. (this volume). Behavioral readiness in this context would reflect the individual or unit's willingness and capability to shift to and enact performance strategies that are fundamentally different from existing routines (cf. Bierman et al., 2009).

While all adaptive performance situations reflect a common performance requisite to fundamentally shift existing task or mission strategies, they may still vary along some important dimensions that influence requirements for adaptive readiness. Pulakos et al. (2000, p. 617) specified eight dimensions of adaptive job performance. These were *handling emergencies or crisis situations*; *handling work stress*; *solving problems creatively*; *dealing with uncertain and unpredictable work situations*; *learning work tasks*, *technologies, and procedures; demonstrating interpersonal adaptability*; *demonstrating cultural adaptability*; and *demonstrating physically oriented adaptability*. These dimensions encode specific tasks and activities that reflect these different forms of adaptive performance. For example, the dimension of demonstrating interpersonal adaptability includes activities such as "working well and developing effective relationships with highly diverse personalities" and "demonstrating keen insight of others' behavior and tailoring own behavior to persuade, influence, or work more effectively with them" (Pulakos et al., 2000, p. 617). Pulakos et al. (2000) provided evidence supporting this taxonomy from personnel in 24 jobs, including many military occupational specialties.

These dimensions suggest that individuals may face a variety of different challenges and performance demands across adaptive performance situations. These demands can be grouped into categories pertaining to how much cognitive, social, and emotional resources they require for effective adaptation. Adaptive situations requiring heavy deployment of cognitive resources can be defined as imposing a high *cognitive load* on the performer, situations requiring high social resources carry a high *social load*, and situations demanding heavy emotional resources have a high *emotional load*. This grouping can be compared to one offered by Mueller-Hanson, White, Dorsey, and Pulakos (2005), who distinguished among mental, interpersonal, and physical adaptability. However, while these categories group different sets of performance tasks under categories of adaptability, they do not necessarily correspond to the different types of psychological load or resources needed to complete these tasks. Performance tasks that would be grouped under one type of adaptability described by Mueller-Hanson et al. (2005) can still carry multiple kinds of loads. For example, handling emergency or crisis situations was listed by Mueller-Hanson et al. under mental adaptability and included the task of "maintaining emotional control and objectivity during emergencies while keeping focused on the situation at hand" (p. A-1). Such tasks will carry high amounts of both cognitive and emotional load; indeed, we suspect that such situations may carry a higher emotional than cognitive load, raising different implications for predictive attributes and training strategies. Likewise, demonstrating interpersonal adaptability can impose high levels of both cognitive and social loads on performers. Further, when having to handle work stressors that are interpersonal in nature (e.g., Fiedler, 1995) and that involve complex organizational problems, individuals are likely to experience high cognitive, social, and emotional loads. Thus, we would argue the delineation of different types of adaptive performance situations can be driven by precise specifications of the extant loads in each context.

5.1.1 Cognitive Load and Adaptive Readiness

Cognitive load theory (Paas & van Merriënboer, 1994; Sweller, 1988; Sweller, van Merriënboer, & Paas, 1998) has been used to describe the "load that performing a particular task imposes on the cognitive system" (Paas & van Merriënboer, 1994, p. 353). It reflects the amount of cognitive capacities and information-processing resources that need to be (or are actually) expended for effective performance (Paas & van Merriënboer, 1994). While cognitive load research has focused almost exclusively on learning and instructional design (Paas, Van Gog, & Sweller, 2010; Sweller, 1988), it has been applied to the development of skills necessary to "dynamically adjust cognitive activities based on flexible knowledge, nonroutinely approach new tasks and ideas, and rapidly acquire as well as use new knowledge and skills in practice" (Kalyuga, Renkl, & Paas, 2010, p. 175). It has also been used to understand coping mechanisms in posttraumatic stress syndrome (Aikins et al., 2009), an example of one of the adaptive performance dimensions presented by Pulakos et al. (2000; i.e., handling stress). Thus, the concept of cognitive load can be easily applied to understanding adaptive readiness and adaptive performance.

Cognitive load has been defined as deriving from both task and person attributes, including their interaction (Paas & van Merriënboer, 1994). According to Paas and van Merriënboer, task factors that can increase cognitive load include novelty, negative consequences for failure, and various external stressors such as time pressure, high noise, and extreme temperatures; person attributes include cognitive abilities, cognitive styles, and existing knowledge stores. Schroder, Driver, and Streuferi (1967) offer a formulation that describes the information attributes that contribute to task complexity. These include information load, or the number of sources requiring focused attention, information diversity, or the variety in information sources, and rate of information change, or the dynamism that characterizes information sources.

By virtue of the need to alter existing performance strategies, and in many cases come up with novel responses, adaptive situations can heighten the information-processing requirements for performers. Zaccaro and his colleagues (Ely et al., 2009; Zaccaro, Banks, Kiechel-Koles, Kemp, & Bader, 2009; see also Burke et al., 2006) specified six problem-solving processes related to adaptation, four of which are explicitly cognitive in nature. These four include (a) scanning operating environments for changes in situational patterns and critical performance requirements, (b) interpreting the meaning of observed environmental changes, (c) formulating adaptive responses to environmental change, and after an adaptive response has been implemented (d) monitoring the situation to determine if successful adaptation has occurred. The first two processes represent components of situational awareness (Endsley, 1997), but in this instance they refer to identification of what is changing in the environment and the interpretation of these changes. Core cognitive capacities necessary to effectively engage these processes include skills in altering one's cognitive frame when scanning the operational environment, making sense of observed changes, and coming up with novel responses (Ely et al., 2009; Nelson et al., 2010). These processing demands heighten the cognitive load in adaptive versus more routine kinds of situations. Thus, the cognitive component of adaptive readiness entails having the potential to employ frame-switching skills when engaged in the aforementioned adaptation processes.

5.1.2 Social Load and Adaptive Readiness

Adaptive performance dimensions can also vary in terms of how much social resources they require of performers. At a simple low level, social load can entail working with people that are known to the performer, enacting common and familiar interpersonal routines. In teams, for example, members who have worked together for a long period, have developed strong shared mental models (Cannon-Bowers, Salas, & Converse, 1993), and are performing routine activities will not likely need to employ significant social resources. However, as situations increase in social complexity, greater amounts of social resources become necessary for operational effectiveness. Social complexity refers to the number and variety of individuals, teams, and organizations that are actors within performance episodes (Zaccaro, 2001). Such variety can be reflected in surface features, such as gender, race, cultural background, and functional expertise, and deep features such as personality, attitudes, and beliefs (Harrison, Price, & Bell, 1998; Harrison, Price, Gavin, & Florey, 2002).

Social resources that are employed as social load increases include both cognitive and behavioral, or interpersonal, activities. Cognitive activities may include social perception and the application of social schemas to interpret social cues (Fiske & Taylor, 1991; Moskowitz, 2005). Such application may entail the use of perspective taking or adopting the frame of reference used by other social actors (Galinsky, Maddux, Gilin, & White, 2008). Such activities may also involve the development of new schemas, or elaboration of existing ones, to apply to novel social situations (Fiske & Taylor, 1991). Thus, in part, social load can overlap with cognitive load when social information-processing demands rise as a function of social complexity.

Greater numbers of social stakeholders, and higher social variety, however, will also likely require a broader array of behavioral and interpersonal responses. Hooijberg (1996) defined such responsiveness as reflecting behavioral complexity. He specified two skill components of behavioral complexity—*behavioral repertoire* and *behavioral differentiation*. Behavioral repertoire refers to the multiplicity of behaviors and roles individuals can enact across different social situations. However, according to Hooijberg, a wide behavioral repertoire is insufficient for successful adaption to social complexity; there is also a need for an ability to determine and apply the most appropriate response to different situational contingencies. This behavioral differentiation is similar to the concept of behavioral flexibility offered by Zaccaro, Gilbert, Thor, and Mumford (1991). These notions suggest that adaptive readiness in high social load situations requires having ready skills in perceiving and understanding complex social environments and in deploying appropriate interpersonal responses.

5.1.3 Emotional Load and Adaptive Readiness

The concept of "emotional load" has not appeared much, if at all, in the human performance literature. The term is more commonly found in the psychophysiology literature and used to describe qualities or elements of an organism's environment (e.g., Adam, Mallan, & Lipp, 2009; Franz, Schaefer, & Schneider, 2003) or requirements of particular tasks (e.g., Thoeringer et al., 2007) that elicit psychophysical emotional responses. In the work literature, while not applying the term emotional load, several researchers have noted how job demands can trigger emotional states that tax employee resources. More specifically, the combination of high job demands, such as time pressure, work overload, role ambiguity, and conflict, and employees' resources to address such demand determines subsequent job stress and job strain (Bakker & Demerouti, 2007). The literature on a related construct "emotional labor" describes "the effort, planning, and control needed to express organizationally desired emotions during interpersonal transactions" (Morris & Feldman, 1996, p. 987; Brotheridge & Grandey, 2002). The labor necessary to expend would rise as a function of decreased congruence between felt emotions and desired ones (Morris & Feldman, 1997). This formulation suggests that emotional labor derives from situational demands. Grandey (2000) described two sets of situational antecedents of emotional labor—interaction expectations and emotional events. Certain jobs (e.g., customer service positions) carry expectations and display rules (Pugliesi, 1999) that incumbents exhibit certain emotions despite felt emotional states. Emotional events can increase the resources necessary for emotional labor, when they increase the contrast between felt emotions and those called for by emotional display rules (Grandey, 2000).

These related themes from disparate literatures support the idea that situations can exert emotional demands on individuals, causing the expenditure of various resources to effectively address them. We would argue that the level of such demands in any situation determines its level of emotional load. Accordingly, we define emotional load as the presence of high levels of emotion-inducing stimuli in the operating environment that require the deployment of mental, social, and emotional coping resources to maintain operational effectiveness. Emotional components of adaptive readiness, then, would pertain to the skills and competencies that foster successful deployment and outcomes of such responses.

5.1.4 Summary

We have suggested in this section that adaptive performance situations can vary according to the cognitive, social, and emotional loads they exert on performers. These loads correspond, respectively, to the degree of cognitive, social, and emotional readiness necessary for operational effectiveness in adaptive situations.

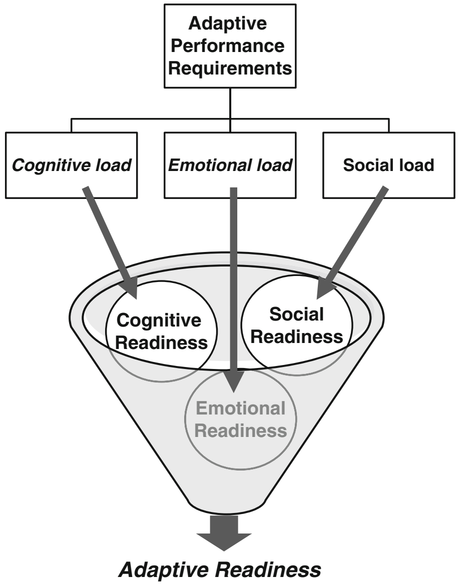


Fig. 5.1 Components of adaptive readiness

Adaptive readiness reflects a combination of these three forms of readiness. Figure 5.1 illustrates this model of adaptive readiness.

We want to make several points about this model. First, we argue that the proportion of cognitive, emotional, and social components of adaptive readiness will likely vary across different adaptive situations, because the load extant in each situation may have a correspondingly different mix of cognitive, emotional, and social elements. Thus, a complex task requiring creative problem solving and learning new technologies in a military school house setting will carry a high cognitive load, but perhaps a low emotional and social load. However, conducting such problem solving with a team composed of domestic and foreign officers heightens the social load in the situation. Finally, requiring such planning in a time-compressed planning phase of a critically important combat mission will greatly add to the emotional load confronting the performers. In each situation, adaptive readiness is necessary, but its components will change significantly according to its load elements.

A second point about our model is that the different components of adaptive readiness are not mutually exclusive. Cognitive skills are at times necessary for operational effectiveness in situations with high social or emotional load. Socially effective behavior often requires the application of social perception and social sense-making processes (Maitlis, 2005; Zaccaro et al., 1991)—i.e., socially oriented cognitive processes—to determine what the most appropriate responses should be when presented with a variety of social cues. Likewise, research on emotional intelligence suggests that the management of emotions entails in part the identification and interpretation of emotions in the self, and in others, as well as regulating emotions in the self and others (Brackett & Mayer, 2003; Caruso, Mayer,

& Salovey, 2002; Mayer & Salovey, 1997). Thus, situations carrying high emotional load may require cognitive competencies to understand emotions and social competencies in managing emotions in interactions between the self and others.

A third point that derives from our model suggests that if elements of adaptive readiness vary according to the blend of cognitive, social, and emotional load in the extant situation, then different mixes of core skills and competencies will be needed across different situations (cf. Ployhart & Bliese, 2006). Moreover, if a situation has high levels of multiple loads, then performers will need high levels of those knowledge skills and abilities that correspond to each form of readiness in the situation. Thus, if a performer has only high cognitive readiness in a situation that also has high social and/or emotional loads, then that individual may be no more prepared for operational effectiveness than performers with lower levels of cognitive readiness. This premise is similar to the pattern approach to leader traits and performance (Foti & Hauenstein, 2007; Zaccaro, 2007), which argues that effective leaders need high levels of cognitive, social, and motivational attributes to lead effectively. The absence of any one set of attributes will result in leadership performance no better than if a person possessed none of the attributes. We believe a similar framework may apply to understanding how the mix of KSAs defines adaptive readiness and its relationships to operational effectiveness in complex domains.

If the adaptive readiness requires different blends of cognitive, social, and emotional competencies, then different training strategies may be necessary to foster growth in adaptive readiness for situations with varying mixes of situational load. Most adaptation training strategies have focused primarily on developing the cognitive skills required to foster adaptive expertise (Bell & Kozlowski, 2002, 2008; Ely et al., 2009; Nelson et al., 2010). We would argue that generic adaptive readiness training will require a focus on growing those competencies for adaptive situations having high cognitive, social, and emotional loads. If trainers are targeting specific adaptive situations, then the preferred training strategy would need to derive from the particular mix of load elements in those situations. In the remaining section of this chapter, we describe in more detail the KSAs that we believe comprise adaptive readiness. We also briefly review the training strategies that are likely to foster growth in the cognitive, social, and emotional components of adaptive readiness.

5.2 Cognitive, Social, and Emotional Competencies that Contribute to Adaptive Readiness: Implications for Readiness Training

We have argued that adaptive readiness reflects an individual's preparation or potential to exhibit a range of personal attributes that are related to effective adaptive performance. While these attributes are likely to include personality and general mental ability (Mueller-Hanson et al., 2005; Pulakos, Dorsey, & White, 2006), we

Table 5.1 Skills, competencies, and instructional strategies that contribute to adaptive readiness

Dimensions of adaptive readiness	Key skills and competencies	Training and development strategies
Cognitive readiness	Self-regulation	Self-regulation training
	Metacognitive thinking	Active learning
	Cognitive flexibility	Error management training
	Frame-changing	Experiential variety
	Creative thinking	Developmental work experiences
	Adaptive expertise	
Social readiness	Social perceptiveness	Developmental work experiences within socially and/or culturally diverse domains
	Perspective taking	
	Behavioral complexity/behavioral flexibility	Cultural assimilators
	Cultural acuity	
Emotional readiness	Emotion regulation and management	Emotion knowledge training
	Emotion identification—self/others	
	Emotion expression	
	Resilience, grit, hardiness	Emotion regulation and management training
	Stress resistance	Stress resistance training
	Stress management	Stress management

are limiting our coverage to those attributes that are likely to be more trainable (i.e., KSAs). Moreover, while we describe KSAs that pertain, for example, to cognitive readiness, we are not covering competencies for such readiness across all kinds of situations. Instead, we are focusing on attributes that foster the cognitive, social, and emotional components of adaptive readiness, that is, readiness for operational effectiveness in situations requiring high levels of adaptation. Finally, we do not pretend that this list is an exhaustive one. However, we believe these to be the most critical KSAs contributing to the three components of adaptive readiness. Table 5.1 indicates (a) each dimension of adaptive readiness; (b) key knowledge, skills, and competencies that compose each form of readiness; and (c) training and development strategies that target growth in some or all of these competencies.

5.2.1 Cognitive Components of Adaptive Readiness

Table 5.1 lists several cognitive competencies that researchers have linked to effective adaptive performance (Mueller-Hanson et al., 2005; Ployhart & Bliese, 2006, Pulakos et al., 2002, 2006; Zaccaro et al., 2009). These include self-regulation, metacognitive thinking, cognitive flexibility and frame-switching, creative thinking, and adaptive expertise. Self-regulation skills refer to competencies in "planning, goal setting, goal monitoring, evaluating goal progress (particularly the detection of discrepancies between ongoing actions and goal progress standards), discrepancy reduction, and goal completion" (Ely et al., 2009, p. 180; see also Karoly, 1993). Such skills are important for adaptation because they are directed to tracking ongoing performance and helping the individual change goal paths in the face of environmental disruptions (Ely et al., 2009; Karoly, 1993). Metacognitive-thinking skills are important because they help performers monitor and regulate how and when they use cognitive processes to throughout stages of adaptive problem solving (Davidson, Deuser, & Sternberg, 1994; Kozlowski, 1998). Bell and Kozlowski (2008) found that metacognitive processes occurring during an adaptive training trial fostered self-evaluation activities and, through such activities, growth in knowledge that contributed to adaptive performance. Skills in both self-regulation and metacognition contribute to adaptive readiness because they prime performers to attend more closely to changing elements of the operating environment and to adjust accordingly to situational understanding and performance strategies.

Cognitive flexibility and, more specifically, frame-changing skills have also been linked to effective adaptation (Griffin & Hesketh, 2003; Nelson et al., 2010). Cognitive flexibility refers to "a person's (a) awareness that in any given situation there are options and alternatives available, (b) willingness to be flexible and adapt to the situation, and (c) self-efficacy or belief that one has the ability to be flexible" (Martin & Anderson, 1998, p. 1). Individuals who adopt a more cognitively flexible approach to problem solving are more likely to explore different cognitive frames when trying to construct and understand the problem space, as well as generate and evaluate potential solutions (Spiro, Feltovich, & Coulson, 1996). Such wide-ranging exploration of the problem space should also facilitate the development of adaptive expertise (Griffin & Hesketh, 2003).

Nelson et al. (2010, p. 133) defined frame-changing skills as "the capacity to switch among various perspectives or frames of reference" at different phases of adaptive problem solving. Thus, individuals can switch among alternate frames when (a) scanning changing operational environments, (b) making sense of environmental changes, and/or (c) deriving adaptive solutions. Horn (2008) defined three components of frame-changing processes—*frame-breaking*, *frame-switching*, and *frame integration*. Frame-breaking reflects skill in recognizing that existing conceptual models can no longer be applied to changing operational environments (DeYoung, Flanders, & Peterson, 2008; London, 1989). Frame-switching entails an exploration of alternate ways adaptive problems can be defined and resolved (Marshall, 1995). Frame integration refers to skill in integrating newly explored cognitive frames into existing cognitive schemas (Jacobs & Jaques, 1987; Jacobs & McGee, 2001). These processes and corresponding skills facilitate growth in adaptive expertise as performers are increasingly able to link different solution frames to different kinds of problems (Zaccaro, 2009).

Frame-changing is a highly effortful and difficult cognitive process to accomplish (Nelson et al., 2010; Zyphur, 2009). Zyphur (2009) noted that individuals who attempt frame-changing needed to "recognize their enacted mindsets and then consciously evaluate and alter them—no easy task" (p. 685). However, because of the central importance of these processes for effective adaption, adaptive readiness entails having the willingness, capacity, and preparation to engage in cognitive frame-changing. For this reason, cognitive readiness for adaptive situations may need to center most strongly on this and related skills.

Mumford and Gustafson (1988) defined creative thinking in part as entailing "processes underlying an individual's capacity to generate new ideas or understandings" (p. 28; also see Hong, this volume). The ability to effectively engage such processes when necessary should contribute significantly to adaptation, especially when changing operational environments pose novel or unusual problems to performers (Pulakos et al., 2006). The use of creative-thinking skills in adaptive domains, combined with use of frame-changing skills, self-regulation, and metacognition, should help performers understand more readily when and how different kinds of solutions will apply to different types of problems. This understanding has been labeled adaptive expertise (Kozlowski, 1998; Smith et al., 1997). Adaptive experts know at a deep, principled level what and how problem and solution constructs are connected and, more importantly, what contextual parameters determine these connections (Kozlowski, 1998). Indeed, Kozlowski (1998) notes, "Adaptive experts are able to recognize changes in task priorities and the need to modify strategies and actions" (p. 119).

5.2.2 Cognitive Training Strategies for Adaptive Readiness

The training strategies necessary to grow adaptive readiness need to be fundamentally different from those used to develop other types of performance skills (Smith et al., 1997). Traditional strategies focus on routinizing the application of such skills. Adaptation training needs to target adaptive expertise and how performers connect contextual parameters, problem elements, and solution strategies (Bell & Kozlowski, 2002, 2008; Ely et al., 2009; Kozlowski et al., 2001). In Table 5.1, we list five strategies that should contribute to the development of the cognitive skills associated with adaptive readiness (cf. Ely et al., 2009; Kozlowski, 1998; Kozlowski et al., 2001; Nelson et al., 2010). These are self-regulation training, active learning, error management training, experiential variety, and developmental work experiences.

Self-regulation training entails instruction in processes such as goal setting, self-monitoring, and self-evaluation to regulation performance progress (Sitzmann, 2007). Trainees are provided prompts during training to engage in such processes until they become more routinized (Ely et al., 2009; Sitzmann, 2007). Bell and Kozlowski (2008) note that self-regulation processes may also follow from using active learning as an instructional strategy. They argue that such approaches, which give the trainees substantial control over their learning process and progress, can be particularly useful in developing adaptive expertise because they help individuals learn to use their existing knowledge to derive solutions to different or novel problems (i.e., adaptive transfer; Ivancic & Hesketh, 2000). Bell and Kozlowski (2002) coupled such approaches with the use of *adaptive guidance*, or a "training strategy that provides trainees with diagnostic and interpretive information that helps them make effective learning decisions" (p. 269). This strategy entails use of tailored information to guide learners in making decisions about what instructional exercises and material will best foster skill growth in an active learning context.

Bell and Kozlowski (2008; see also Keith and Frese, 2005, 2008) also argue that training strategies using error management or exploratory learning approaches are effective means of developing adaptation skills. In these approaches, learners are encouraged to explore an unfamiliar problem space and instructed that errors are not only acceptable, but critical to the learning process. Researchers have linked emotion management training and exploratory learning to the development of adaptive expertise and adaptive transfer (Bell & Kozlowski, 2008; Ely et al., 2009). Nelson et al. (2010) noted that in these and other forms of active learning, instructional strategies should include the use of experiential variety or the exposure of learners to "stimuli or practice scenarios in training that vary in either surface or structural details that in turn require changes to performance strategies" (p. 133). They argued that when learners encountered qualitatively different problem scenarios during practice trials, they become more adept at changing the cognitive frames used to interpret, understand, and solve different types of problems.

A central principle underlying all of these adaptive training strategies is having individuals experiencing new kinds of problems as part of the instructional strategy. Most of the studies supporting this approach have been conducted in formal training contexts. Zaccaro and Banks (2004) argued that developmental work experiences, or stretch assignments (Ohlott, 2004) encountered in one's job context, can be effective tools for developing adaptation skills. These kinds of experiences entail giving to job incumbents assignments that challenge their current skill sets and cognitive frames. Both Banks (2006) and Horn (2008) found that developmental work experiences were associated with indicators of adaptive performance.

5.2.3 Social Components of Adaptive Readiness

The social elements of adaptive readiness entail being prepared to adjust to working with and across different types of people and social groups, including those from different cultures. The skills and competencies associated with such adaptation have been grouped under social and cultural intelligence (Earley & Ang, 2003; Pulakos et al., 2006; Zaccaro et al., 1991). Zaccaro et al. (1991) defined social intelligence as reflecting capacities to engage in effective social perception and awareness, as well behavioral flexibility, or the capacity to respond appropriately across different situations. Hooijberg (1996) elaborated the latter competency as behavioral complexity, which involves the possession of a broad behavioral repertoire and the capacity to perform behaviors in their repertoire in adaptive ways according to situational requirements.

Cultural acuity refers to the extension of social perceptions skills to understanding persons and social dynamics from multiple cultures (see similar themes in Earley & Ang, 2003). Such acuity has two foci—the self and the team (Chiara et al., 2010). According to Chiara et al. (2010), self-focused cultural acuity refers to understanding one's own culture-related biases and how they might influence interactions with individuals from other cultures. This skill reflects cultural self-awareness (Earley & Ang, 2003). A team-focused acuity refers to understanding how culture will affect interaction dynamics within a team that (a) is embedded within another culture or (b) is composed of members from different cultures. Sutton, Pierce, Burke, and Salas (2006) extended these ideas in their notion of cultural adaptability, defined as "the ability to understand one's own and others' cognitive biases and to adapt, as necessary, to ensure successful team performance" (p. 144). Their notion adds skill in adjusting behavioral responses to cultural variants.

Social intelligence and cultural acuity can often entail trying to be aware of and appreciate the understanding other people have of a particular social context. This awareness refers to social perspective taking (Johnson, 1975; Roan et al., 2009), defined as

> Taking the perspective of another person is the ability to understand how a situation appears to another person and how that person is reacting cognitively and emotionally to the situation. It is the ability to put oneself in the place of others and recognize that other individuals may have points of view different from one's own. (Johnson, 1975; p. 241)

Such perspective taking can foster adaptability because it facilitates the likelihood that performers will adopt the most appropriate strategy or behavioral response in socially diverse contexts. The selection of socially appropriate responses derives from a clear and deep understanding of how such responses are likely to affect others in the context; such understanding comes more readily to those individuals that can consider responses from the perspective of other who will be their recipients (Roan et al., 2009).

5.2.4 Social Training Strategies for Adaptive Readiness

The training and development of social competencies related to adaptive readiness entail having learners experience a diversity or variety of social contexts, with a focus on (a) understanding differences across such contexts and (b) learning context-specific and appropriate social behaviors. Such learning can occur through the use of developmental work experiences that require individuals to work across different social contexts (Ohlott, 2004). With respect to leader development, for example, Ohlott (2004, p. 161) recommends that leaders be given assignments to lead "people who are not like themselves;" doing so would challenge them "to move beyond their own beliefs and perspectives to understand personal, business, and workplace issues from perspectives that may differ greatly from, and sometimes even conflict with, their own." Such experiences would not only obviously facilitate skill in perspective taking but also contribute greatly to the social knowledge structures that contribute to effective asocial intelligence (Zaccaro et al., 1991).

The development of cultural acuity entails the same principle of having individuals experience social diversity, except now across cultural boundaries. Being immersed into foreign cultures can create those "mind-altering, head-cracking experiences" (Gregersen, Morrison, & Black, 1998, p. 30) that foster the development of self-knowledge necessary for effective self-focused cultural acuity, as well as better understanding of the different cultural variants in behavior necessary for cultural adaptability. Indeed, regarding the development of such skills in leaders, Nelson et al. (2010) noted that "leaders would need (a) learning experiences that help them discover new, culturally variant leadership frames, and (b) guidance on the appropriate application of these frames" (p. 139). Not all such experiences need to occur in situ within foreign cultures—Bhawuk (2001) recommends the use of *cultural assimilators* or as defined by Nelson et al. (2010), "scenario-based, feedback-rich exercises that can provide (a) intensive culture-specific information to prepare leaders to adapt to specific cultural contexts, or (b) broad, culture-general theory, to help them focus on cultural dimensions that apply to many cultures" (p. 140; see Abbe, Gulick, & Herman, 2007 for a relevant review). Note that such exercises can be incorporated into formal training exercises designed to grow cultural adaptability; Nelson et al. argue for the incorporation of experiential variety into such exercises to make them even more effective.

5.2.5 Emotion Components of Adaptive Readiness

When situations requiring adaptation carry a high emotional load, adaptive readiness includes the potential to use both emotion understanding and management skills. Such readiness may also need to include capacities to withstand and work effectively under highly stressful circumstances. Emotion understanding and management skills are defined as components of emotional intelligence (Caruso et al., 2002; Mayer & Salovey, 1997). Mayer and Salovey (1997; see also Caruso et al., 2002) defined four competencies that contribute to the emotional intelligence—(1) the accurate identification of emotions and feelings, (2) interpretation and accurate understanding of emotions, (3) the effective use of emotions in social problem solving, and (4) the management and control of one's own emotions within the context of problem solving. In adaptive contexts that are high in emotion load, operational effectiveness will often require emotion management strategies before one can utilize problem-solving processes—performers need to understand and control their own emotions before they can begin to think effectively about an adaptive solution. Also the derivation and implementation of adaptive solutions in such contexts may also call for performers to help their colleagues and teammates to manage their own emotional reactions. There has been little if any research that has provided empirical evidence linking these skills to adaptation. We encourage such research, suspecting that it will endorse their validity for predicting adaptive performance.

The understanding and management of emotions is one aspect of adaptive readiness. We expect that adaptation in many highly stressful and emotion-laden situations requires a degree of grit, mental toughness, and hardiness that helps the performer persist through difficult and challenging circumstances (Duckworth, Peterson, Matthews, & Kelly, 2007; Loehr, 1986; Maddi, 2007; Mueller-Hanson et al., 2005). In essence, these attributes reflect ability to remain calm and composed even under dire or very stressful circumstances (Loehr, 1986). Studies by Bartone (2000, 2006) show that these qualities are related to operational effectiveness under combat situations, one of Pulakos et al.'s (2000) adaptive performance dimensions.

5.2.6 Emotion Training Strategies for Adaptive Readiness

Caruso and Wolfe (2004) argued that individuals could indeed be trained in emotional intelligence skills. Such training would consist of formal instruction on the nature of emotions and understanding their role in behavior. It would also include practice in the regulation of emotions and particularly in how to "apply specific emotions in … everyday life" (Nelis, Quoidbach, Mikolajczak, & Hansenne, 2009, p. 37). Using this combined instructional strategy, Nelis et al. (2009) produced an increase in emotion identification and management skills that persisted 6 months after training. Similar findings were reported by Groves, McEnrue, and Shen (2008). Clarke (2006, 2010) described a training strategy that successfully used work team situations to explore and examine emotional knowledge and to practice emotion regulation situations in ongoing work assignments. Taken together, these studies suggest that emotion identification, understanding, and regulation skills that foster adaptive readiness can be developed through targeted formal and on-the-job instructional strategies.

Attributes as grit, resilience, and hardiness have often been described as dispositional qualities of the individual and therefore as not easily trainable (e.g., Mueller-Hanson et al., 2005). Maddi and his colleagues (2007; Maddi, Kahn, & Maddi, 1998), though, have demonstrated some success in fostering hardiness in adults. Moreover, researchers have suggested that controlled exposure to high-stress training conditions can be effective in developing resilience. For example, Paton (2006) argued that training simulations for police officers should reflect the conditions they might face in their dangerous operating environment. He noted that such exposure in training can "help increase knowledge of stress reactions and provide opportunities for officers to rehearse strategies to deal with them" (p. 3). The US Army has used variations of such extreme stressor exposure as part of their survival training courses (e.g., Morgan et al., 2000). Finally worksite stress management training programs, including stress inoculation programs (Saunders, Driskell, Johnston, & Salas, 1996), have been successful in helping workers develop cognitive and behavioral strategies that foster resilient reactions to work stressors (Richardson &

Rothstein, 2008). Thus, we believe that a combination of stress management training with exposure to extreme stressors in training scenarios can foster adaptive readiness for situations carrying high emotion loads.

5.3 Summary

In this chapter, we have made several key points about readiness for operational effectiveness in environments requiring adaptation. First, adaptive performance situations will vary in terms of their cognitive, social, and emotional loads. Some situations may carry high levels of all three types of performance requirements. Second, success in such situations will require varying degrees of not just cognitive readiness but social and emotional readiness as well. The literature on human performance in complex problem domains has emphasized primarily cognitive readiness; we suggest an expansion to other forms as well. Third, cognitive, social, and emotional elements of adaptive readiness reflect different sets of KSAs and competencies. Overall adaptive readiness will depend upon the combination of KSAs that corresponds to the load mix in a particular situation. Accordingly, in a situation with high cognitive social and emotional load, cognitive readiness will not be enough to ensure overall operational effectiveness. Finally, different training strategies will be needed to foster cognitive, social, and emotional readiness. And, again, when the situation load mix reflects high levels of more than just cognitive load, adaptive training strategies will have to focus on measurement and assessment of this mix as well as a broader range of cognitive and noncognitive skills than they do in the present. We expect that such a focus will foster a greater all around readiness to adapt successfully in multiple kinds of complex problem domains.

Acknowledgments The work reported herein was partially supported by a grant from the Office of Naval Research, Award Number N000140810126. The findings and opinions expressed here do not necessarily reflect the positions or policies of the Office of Naval Research.

References

Abbe, A., Gulick, L. M. V., & Herman, J. L. (2007). *Developing cross-cultural competence in military leaders: A conceptual and empirical foundation*. Arlington, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Adam, A. R., Mallan, K. M., & Lipp, O. V. (2009). The effect of emotional and attentional load on attentional startle modulation. *International Journal of Psychophysiology, 74*, 266–273.

Aikins, D. E., Johnson, D. C., Borelli, J. L., Klemanski, D. H., Morrissey, P. M., Benham, T. L., et al. (2009). Thought suppression failures in combat PTSD: A cognitive load hypothesis. *Behaviour Research and Therapy, 47*, 744–751.

Armenakis, A., Harris, S. G., & Mossholder, K. W. (1993). Creating readiness for organizational change. *Human Relations, 46*, 681–703.

Bakker, A. B., & Demerouti, E. (2007). The job demands resources model: State of the art. *Journal of Managerial Psychology, 22*, 309–328.
Banks, D. (2006). *Stretch experiences and leader development: The relationships among work experience, individual differences, contextual variables, and leader adaptability*. Unpublished dissertation, George Mason University.
Banks, D., Bader, P. K., Fleming, P. J., Zaccaro, S. J., & Barber, H. F. (2001). *Leader adaptability: The role of work experiences and individual differences.* Paper presented at the 16th Annual Conference of the Society for Industrial and Organizational Psychology, San Diego, CA.
Bartone, P. T. (2000). Hardiness as a resiliency factor for United States forces in the Gulf War. In J. M. Violanti, D. Paton, & C. Dunning (Eds.), *Posttraumatic stress intervention: Challenges, issues, and perspectives* (pp. 115–133). Springfield, IL: Thomas.
Bartone, P. T. (2006). Resilience under military operational stress: Can leaders influence hardiness? *Military Psychology, 18*, S131–S148.
Bell, B. S., & Kozlowski, S. W. J. (2002). Adaptive guidance: Enhancing self-regulation, knowledge, and performance in technology-based training. *Personnel Psychology, 55*, 267–307.
Bell, B. S., & Kozlowski, S. W. J. (2008). Active learning: Effects of core training design elements on self-regulatory processes, learning, and adaptability. *Journal of Applied Psychology, 93*, 296–316.
Bhawuk, D. P. S. (2001). Evolution of culture assimilators: Toward theory-based assimilators. *International Journal of Intercultural Relations, 25*, 141–163.
Bierman, K. L., Torres, M. M., Domitrovich, C. E., Welsh, J. A., & Gest, S. D. (2009). Behavioral and cognitive readiness for school: Cross-domain associations for children attending Head Start. *Social Development, 18*, 305–323.
Brackett, M. A., & Mayer, J. D. (2003). Convergent, discriminant, and incremental validity of competing measures of emotional intelligence. *Personality and Social Psychology Bulletin, 29*, 1147–1158.
Brotheridge, C. M., & Grandey, A. A. (2002). Emotional labor and burnout: Comparing two perspectives of "People Work". *Journal of Vocational Behavior, 60*, 17–39.
Burke, C. S., Stagl, K. C., Salas, E., Pierce, L., & Kendall, D. (2006). Understanding team adaptation: A conceptual analysis and model. *Journal of Applied Psychology, 91*, 1189–1207.
Cannon-Bowers, J. A., Salas, E., & Converse, S. (1993). Shared mental models in expert team decision making. In N. J. Castellan Jr. (Ed.), *Current issues in individual and group decision making* (pp. 221–246). Hillsdale, NJ: Erlbaum.
Caruso, D. R., Mayer, J. D., & Salovey, P. (2002). Emotional intelligence and emotional leadership. In R. Riggio (Ed.), *Multiple Intelligences and Leadership*. Mahwah, NJ: Erlbaum.
Caruso, D. R., & Wolfe, C. J. (2004). Emotional intelligence and leadership development. In D. Day, S. J. Zaccaro, & S. Halpin (Eds.), *Leader development for transforming organizations*. Mahwah, NJ: Erlbaum.
Chan, D. (2000). Understanding adaptation to changes in the work environment: Integrating individual difference and learning perspectives. *Research in Organizational Behavior, 18*, 1–42.
Chiara, J., Burke, C. S., Salas, E., Priest-Walker, H., Zaccaro, S., Riedel, S., et al. (2010). *The Multicultural Leadership Assessment and Development (Multi-LeAD) program: A training approach to the development of multicultural leadership skills (ARI Technical Report)*. Arlington, VA: U.S. Army Research Institute for Behavioral and Social Sciences.
Clarke, N. (2006). Developing emotional intelligence through workplace learning: Findings from a case study in healthcare. *Human Resource Development International, 9*, 447–465.
Clarke, N. (2010). Developing emotional intelligence abilities through team-based learning. *Human Resource Development Quarterly, 21*, 119–138.
Cosenzo, K. A., Fatkin, L. T., & Patton, D. J. (2007). Ready or not: Enhancing operational effectiveness through the use of readiness measures. *Aviation, Space, and Environmental Medicine, 78*, B96–B106.
Davidson, J. E., Deuser, R., & Sternberg, R. J. (1994). The role of metacognition in problem solving. In J. Metcalf & A. P. Shimamura (Eds.), *Metacognition: Knowing about knowing*. Cambridge, MA: MIT Press.

DeYoung, C. G., Flanders, J. L., & Peterson, J. B. (2008). Cognitive abilities involved in insight problem solving: An individual differences model. *Creativity Research Journal, 20*, 278–290.
Duckworth, A. L., Peterson, C., Matthews, M. D., & Kelly, D. R. (2007). Grit, perseverance and passion for long term goals. *Journal of Personality and Social Psychology, 92*, 1087–1101.
Earley, P. C., & Ang, S. (2003). *Cultural intelligence: Individual interactions across cultures*. Stanford, CA: Stanford Business Books.
Eby, L. T., Adams, D. M., Russell, J. E. A., & Gaby, S. H. (2000). Perceptions of organizational readiness for change: Factors related to employees' reactions to the implementation of team-based selling. *Human Relations, 53*, 419–442.
Ely, K., Zaccaro, S. J., & Conjar, E. A. (2009). Leadership development: Training strategies for growing adaptability in leaders. In C. Cooper & R. Burke (Eds.), *The peak performing organization* (pp. 175–196). London: Routledge.
Endsley, M. R. (1997). The role of situation awareness in naturalistic decision making. In C. E. Zsambok & G. Klein (Eds.), *Naturalistic decision making* (pp. 269–283). Mahwah, NJ: Erlbaum.
Fiedler, F. E. (1995). Cognitive resources and leadership performance. *Applied Psychology, 44*, 5–28.
Fiske, S. T., & Taylor, S. E. (1991). *Social cognition* (2nd ed.). New York: McGraw-Hill.
Foti, R. J., & Hauenstein, N. M. A. (2007). Pattern and variable approaches in leadership emergence and effectiveness. *Journal of Applied Psychology, 92*, 347–355.
Franz, M., Schaefer, R., & Schneider, C. (2003). Psychophysiological response patterns of high and low alexithymics under mental and emotional load conditions. *Journal of Psychophysiology, 17*, 203–213.
Galinsky, A. D., Maddux, W. W., Gilin, D., & White, J. B. (2008). Why it pays to get inside the head of your opponent: The differential effects of perspective taking and empathy in negotiations. *Psychological Science, 19*, 378–384.
Grandey, A. A. (2000). Emotion regulation in the workplace: A new way to conceptualize emotional labor. *Journal of Occupational Health Psychology, 5*, 95–110.
Gregersen, H. B., Morrison, A. J., & Black, J. S. (1998). Developing leaders for the global frontier. *Sloan Management Review, 40*, 21–32.
Griffin, B., & Hesketh, B. (2003). Adaptable behaviours for successful work and career adjustment. *Australian Journal of Psychology, 55*(2), 65–73.
Groves, K. S., McEnrue, M. P., & Shen, W. (2008). Developing and measuring the emotional intelligence of leaders. *The Journal of Management Development, 27*, 225–250.
Harrison, D. A., Price, K. H., & Bell, M. P. (1998). Beyond relational demography: Time and the effects of surface-and deep-level diversity on work group cohesion. *Academy of Management Journal, 41*, 96–107.
Harrison, D. A., Price, K. H., Gavin, J. H., & Florey, A. T. (2002). Time, teams, and task performance: Changing effects of surface- and deep-level diversity on group functioning. *Academy of Management Journal, 45*(5), 1029–1045.
Holyoak, K. J. (1991). Symbolic connectionism: Toward third-generation theories of expertise. In K. A. Ericsson & J. Smith (Eds.), *Toward a general theory of expertise* (pp. 301–336). Cambridge, England: Cambridge University Press.
Hooijberg, R. (1996). A multidirectional approach toward leadership: An extension of the concept of behavioral complexity. *Human Relations, 49*(7), 917–946.
Horn, Z. N. J. (2008). Linking developmental stretch assignments with adaptive outcomes: The importance of cognitive frame complexity. *Dissertation Abstracts International Section B: Sciences and Engineering, 69*(06), 3867.
Ivancic, K., & Hesketh, B. (2000). Learning from errors in a driving simulation: Effects on driving skill and self-confidence. *Ergonomics, 43*, 1966–1984.
Jacobs, T. O., & Jaques, E. (1987). Leadership in complex systems. In J. Zeidner (Ed.), *Human productivity enhancement*. New York: Praeger.
Jacobs, T. O., & McGee, M. L. (2001). Competitive advantage: Conceptual imperatives for executives. In S. J. Zaccaro & R. Klimoski (Eds.), *The nature of organizational leadership: Understanding the performance imperatives confronting today's leaders* (pp. 42–78). San Francisco: Jossey-Bass.

Johnson, D. W. (1975). Cooperativeness and social perspective taking. *Journal of Personality and Social Psychology, 31*, 241–244.

Kalyuga, S., Renkl, A., & Paas, F. (2010). Facilitating flexible problem solving: A cognitive load perspective. *Educational Psychology Review, 22*, 175–186.

Karoly, P. (1993). Mechanisms of self-regulation: A systems view. *Annual Review of Psychology, 44*, 23–52.

Keith, N., & Frese, M. (2005). Self-regulation in error management training: Emotion control and metacognition as mediators of performance effects. *Journal of Applied Psychology, 90*, 677–691.

Keith, N., & Frese, M. (2008). Effectiveness of error management training: A meta-analysis. *Journal of Applied Psychology, 93*, 59–69.

Kozlowski, S. W. J. (1998). Training and developing adaptive teams: Theory, principles, and research. In J. A. Cannon-Bowers & E. Salas (Eds.), *Decision making under stress: Implications for training and simulation* (pp. 115–153). Washington, DC: American Psychological Association.

Kozlowski, S. W. J., Gully, S. M., Nason, E. R., & Smith, E. M. (1999). Developing adaptive teams: A theory of compilation and performance across levels and time. In D. R. Ilgen & E. D. Pulakos (Eds.), *The changing nature of performance* (pp. 240–292). San Francisco, CA: Jossey-Bass.

Kozlowski, S. W. J., Toney, R. J., Mullins, M. E., Weissbein, D. A., Brown, K. G., & Bell, B. S. (2001). Developing adaptability: A theory for the design of integrated-embedded training systems. In E. Salas (Ed.), *Advances in human performance and cognitive engineering research* (Vol. 1, pp. 59–123). New York: Elsevier Science/JAI Press.

Le, H., Casillas, A., Robbins, S. B., & Langley, R. (2005). Motivational and skills, social, and self-management predictors of college outcomes: Constructing the Student Readiness Inventory. *Educational and Psychological Measurement, 65*, 482–508.

LePine, J. A. (2003). Team adaptation and postchange performance: Effects of team composition in terms of members' cognitive ability and personality. *Journal of Applied Psychology, 88*, 27–39.

LePine, J. A., Colquitt, J. A., & Erez, A. (2000). Adaptability to changing task contexts: Effects of general cognitive ability, conscientiousness, and openness to experience. *Personnel Psychology, 53*, 563–593.

Loehr, J. E. (1986). *Mental toughness training for sports: Achieving athletic excellence*. Lexington, MA: Stephen Greene Press.

London, M. (1989). *Managing the training enterprise: High-quality, cost-effective employee training in organizations*. San Francisco, CA: Jossey-Bass.

Maddi, S. R. (2007). Relevance of hardiness assessment and training to the military context. *Military Psychology, 19*, 61–70.

Maddi, S. R., Kahn, S., & Maddi, K. L. (1998). The effectiveness of hardiness training. *Consulting Psychology Journal: Practice and Research, 50*, 78–86.

Maitlis, S. (2005). The social process of organizational sensemaking. *Academy of Management Journal, 48*(1), 21–49.

Marshall, S. P. (1995). *Schemas in problem solving*. Cambridge, England: Cambridge University Press.

Martin, M. M., & Anderson, A. M. (1998). The cognitive flexibility scale: Three validity studies. *Communication Reports, 11*, 1–9.

Mayer, J. D., & Salovey, P. (1997). What is emotional intelligence? In P. Salovey & D. Sluyter (Eds.), *Emotional development and emotional intelligence: Implications for educators* (pp. 3–31). New York: Basic Books.

Morgan, C. A., III, Wang, S., Mason, J., Southwick, S. M., Fox, P., Hazlett, G., et al. (2000). Hormone profiles in humans experiencing military survival training. *Biological Psychiatry, 47*, 891–901.

Morris, J. A., & Feldman, D. C. (1996). The dimensions, antecedents, and consequences of emotional labor. *Academy of Management Review, 21*, 986–1010.

Morris, J. A., & Feldman, D. C. (1997). Managing emotions in the workplace. *Journal of Managerial Issues, 9*(3), 257–274.

Morrison, J. E., & Fletcher, J. D. (2002). *Cognitive readiness (IDA Paper P-3735).* Alexandria, VA: Institute for Defense Analyses.
Moskowitz, G. (2005). *Social cognition*. New York, NY: Guilford Press.
Mueller-Hanson, R., White, S. S., Dorsey, D. W., & Pulakos, E. D. (2005). *Training adaptable leaders: Lessons from research and practice (Research Report 1844).* Arlington, VA: U.S. Army Research Institute for Behavioral and Social Sciences.
Mumford, M. D., & Gustafson, S. B. (1988). Creativity syndrome: Integrations, application, and innovation. *Psychological Bulletin, 103*, 27–43.
Nelis, D., Quoidbach, J., Mikolajczak, M., & Hansenne, M. (2009). Increasing emotional intelligence: (How) is it possible? *Personality and Individual Differences, 47*, 36–41.
Nelson, J. K., Zaccaro, S. J., & Herman, J. L. (2010). Strategic information provision and experiential variety as tools for developing adaptive leadership skills. *Consulting Psychology Journal: Practice and Research, 62*, 131–142.
O'Neil, H. F. Jr., Allred, K., & Baker, E. L. (1992). *Measurement of workforce readiness: Review of theoretical frameworks* (CSE Tech. Rep. No. 343). Los Angeles, CA: National Center for Research in Evaluation, Standards, and Student Testing, University of California.
Ohlott, P. J. (2004). Job assignments. In C. D. McCauley & E. Van Velsor (Eds.), *The Center for Creative Leadership handbook of leadership development* (2nd ed., pp. 151–182). San Francisco, CA: Wiley.
Paas, F., van Gog, T., & Sweller, J. (2010). Cognitive load theory: New conceptualizations, specifications, and integrated research perspectives. *Educational Psychology Review, 22*, 115–121.
Paas, F., & van Merriënboer, J. J. G. (1994). Instructional control of cognitive load in the training of complex cognitive tasks. *Educational Psychology Review, 6*, 351–372.
Paton, D. (2006). Critical incident stress risk in police officers: Managing resilience and vulnerability. *Traumatology, 23*, 1–9.
Peterson, C. H., Casillas, A., & Robbins, S. B. (2006). The Student Readiness Inventory and the Big Five: Examining social desirability and college academic performance. *Personality and Individual Differences, 41*, 663–673.
Ployhart, R. E., & Bliese, P. D. (2006). Individual adaptability (I-ADAPT) theory: Conceptualizing the antecedents, consequences, and measurement of individual differences in adaptability. In C. S. Burke, L. G. Pierce, & E. Salas (Eds.), *Advances in human performance and cognitive engineering: Vol. 6. Understanding adaptability: A prerequisite for effective performance within complex environments* (pp. 3–39). Amsterdam: Elsevier.
Pugliesi, K. (1999). The consequences of emotional labor: Effects on work stress, job satisfaction, and well-being. *Motivation and Emotion, 23*(2), 125–154.
Pulakos, E. D., Arad, S., Donovan, M. A., & Plamondon, K. E. (2000). Adaptability in the workplace: Development of a taxonomy of adaptive performance. *Journal of Applied Psychology, 85*, 612–624.
Pulakos, E. D., Dorsey, D. W., & White, S. S. (2006). Adaptability in the workplace: Selecting an adaptive workforce. In C. S. Burke, L. G. Pierce, & E. Salas (Eds.), *Advances in human performance and cognitive engineering: Vol. 6. A prerequisite for effective performance within complex environments* (pp. 41–71). Amsterdam: Elsevier.
Pulakos, E. D., Schmitt, N., Dorsey, D. W., Arad, S., Hedge, J. W., & Borman, W. C. (2002). Predicting adaptive performance: Further tests of a model of adaptability. *Human Performance, 15*(4), 299–323.
Richardson, K. M., & Rothstein, H. R. (2008). Effects of occupational stress management intervention programs: A meta-analysis. *Journal of Occupational Health Psychology, 13*, 69–93.
Roan, L., Strong, B., Foss, P., Yager, M., Gehlback, H., & Metcalf, K. (2009). *Social perspective taking (ARI Technical Report 1259).* Arlington, VA: U.S. Army Research Institute for Behavioral and Social Sciences.
Robbins, S., Lauver, K., Le, H., Davis, D., Langley, R., & Carlstrom, A. (2004). Do psychosocial and study skill factors predict college outcomes? A meta-analysis. *Psychological Bulletin, 130*, 261–288.
Saunders, T., Driskell, J. E., Johnston, J. H., & Salas, E. (1996). The effect of stress inoculation training on anxiety and performance. *Journal of Occupational Health Psychology, 1*, 170–186.

Schroder, H., Driver, M., & Streuferi, S. (1967). *Human information processing*. New York: Holt, Rinehart and Winston.

Sitzmann, T. (2007). *Prompting self-regulation to improve learning outcomes in Web-based training*. Presentation at the 22nd Annual Conference of the Society for Industrial-Organizational Psychology Conference, New York, NY.

Smith, E. M., Ford, J., & Kozlowski, S. W. J. (1997). Building adaptive expertise: Implications for training design strategies. In M. A. Quinones & A. Ehrenstein (Eds.), *Training for a rapidly changing workplace: Applications of psychological research* (pp. 89–118). Washington, DC: American Psychological Association.

Smyth, C. C. (2007). Sensitivity of subjective questionnaires to cognitive loading while driving with navigation aids: A pilot study. *Aviation, Space, and Environmental Medicine, 78*, B39–B50.

Spiro, R. J., Feltovich, P. J., & Coulson, R. L. (1996). Two epistemic world-views: Prefigurative schemas and leaning in complex domains. *Applied Cognitive Psychology, 10*, S51–S61.

Sutton, J. L., Pierce, L. G., Burke, C. S., & Salas, E. (2006). Cultural adaptability. In C. S. Burke, L. G. Pierce, & E. Salas (Eds.), *Advances in human performance and cognitive engineering: Vol. 6. A prerequisite for effective performance within complex environments* (pp. 143–173). Amsterdam: Elsevier.

Sweller, J. (1988). Cognitive load during problem solving: Effects on learning. *Cognitive Science, 12*(1), 257–285.

Sweller, J., van Merriënboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*, 251–296.

Thoeringer, C. K., Sillaber, I., Roedel, A., Erhardt, A., Mueller, M. B., Ohl, F., et al. (2007). The temporal dynamics of intrahippocampal corticosterone in response to stress-related stimuli with different emotional and physical load: An in vivo microdialysis study in C57BL/6 and DBA/2 inbred mice. *Psychoneuroendocrinology, 32*, 746–757.

Wesley, P. W., & Buysse, V. (2003). Making meaning of school readiness in schools and communities. *Early Childhood Research Quarterly, 18*, 351–375.

White, S. S., Mueller-Hanson, R., Dorsey, D. W., Pulakos, E. D., Wisecarver, M. M., Deagle, E. A., et al. (2005). *Developing adaptive proficiency in Special Forces officers (Research Report 1831)*. Arlington, VA: U.S. Army Research Institute for Behavioral and Social Sciences.

Zaccaro, S. J. (2001). *The nature of executive leadership: A conceptual and empirical analysis of success*. Washington, DC: APA Books.

Zaccaro, S. J. (2007). Trait-based perspectives in leadership. *American Psychologist, 62*, 6–16.

Zaccaro, S. J. (2009). *Principles and strategies for developing skills that promote adaptive performance*. Arlington, VA: U.S. Army Research Institute for Behavioral and Social Sciences.

Zaccaro, S. J., & Banks, D. (2004). Leader visioning and adaptability: Bridging the gap between research and practice on developing the ability to manage change. *Human Resource Management, 43*(4), 367.

Zaccaro, S. J., Banks, D., Kiechel-Koles, L., Kemp, C., & Bader, P. (2009). *Leader and team adaptation: The influences and development of key attributes and processes (Tech. Rep. No. 1256)*. Arlington, VA: U.S. Army Research Institute for Behavioral and Social Sciences.

Zaccaro, S. J., Gilbert, J. A., Thor, K. K., & Mumford, M. D. (1991). Leadership and social intelligence: Linking social perceptiveness to behavioral flexibility. *The Leadership Quarterly, 2*, 317–347.

Zyphur, M. J. (2009). When mindsets collide: Switching analytical mindsets to advance organization science. *Academy of Management Review, 34*(4), 677–688.

Chapter 6
The Development and Assessment of Cognitive Readiness: Lessons Learned from K-12 Education

Joan L. Herman

6.1 Introduction

This chapter considers what lessons can be learned from the K-12 education sector's experience in defining and assessing readiness. The chapter shares examples of readiness definitions and assessment strategies at three key transition points:

- Are you ready for kindergarten?
- Are you ready for college?
- Are you ready for challenging work?

Just as the concept of cognitive readiness in the military can be conceptualized as readiness to respond to the challenges of what lies ahead—i.e., combat applications and adaptation to unpredictable circumstances (Fletcher, 2004)—so too do students and job applicants need to be ready for the demands of what lies ahead as they transition from one educational level or position to another. Treated in turn are readiness for kindergarten, readiness for college, and readiness for a demanding career. While core knowledge, skills, and dispositions needed for success are categorized differently and the specifics of terminology vary across the three contexts, the chapter uses these examples to consider common characteristics of readiness both historically and currently. These common threads include such capacities as relevant content knowledge; cognitive strategies such as problem solving and analytic reasoning; social competence, including teamwork and leadership; communication; motivation and persistence; and metacognition. The consequences of a mismatch between characteristics of readiness and measures of it also are considered. Finally, the chapter suggests the core elements of training and assessment systems to support readiness.

J.L. Herman (✉)
National Center for Research on Evaluation, Standards, and Student Testing (CRESST), UCLA Graduate School of Education and Information Studies, 311 GSE & IS Building, 300 Charles E. Young Drive North, Los Angeles, CA 90095-1522, USA
e-mail: herman@cse.ucla.edu

H.F. O'Neil et al. (eds.), *Teaching and Measuring Cognitive Readiness*,
DOI 10.1007/978-1-4614-7579-8_6, © Springer Science+Business Media New York 2014

6.2 Readiness for Kindergarten

What does it mean to be ready for kindergarten, a key, early transition for most children, as they move from home and more informal learning environments to public school and formal educational expectations? In times past, being ready for kindergarten has meant reaching age 5 by a particular month in the academic year. No more.

The idea of needing more than to simply reach a particular chronological age to be ready for kindergarten was fueled as part of the national response to *A Nation at Risk*, a prominent national commission report that documented the mediocrity of public education in the United States, the continuing achievement gaps it enabled to continue, and the disappointing performance of American students in international comparisons (National Commission on Excellence in Education, 1983). In response, the Congress established the National Education Goals Panel that was charged with assessing and reporting on state and national progress toward achieving National Education Goals to remedy this situation. First among these goals was that by the year 2000, "All Children in America will start school ready to learn." The goal was an attempt to ameliorate the major existing achievement gaps among students at school entry based on research documenting that children from economically disadvantaged communities started kindergarten with significantly smaller vocabularies and lower cognitive skills than their more advantaged peers (Farkas & Beron, 2004; Hart & Risley, 1995; Lee & Burkham, 2002). Research also supports the strong relationship between these variables and subsequent literacy development and school success (Biemiller, 2006).

Standards and assessment. States and a number of professional organizations responded to this goal by establishing standards for early childhood education, ostensibly bringing together the expertise of researchers and practitioners. These efforts have stimulated considerable public discussion of what it means to be ready for school and the implications of such readiness for the development of early childhood programs and the assessment of young children (see, e.g., Snow & Van Hemel, 2008). While there are various lists of expected competences for school readiness, they all basically boil down to a common set of knowledge, skills, and predispositions that largely echo, albeit in age-appropriate definitions, the categories comprising cognitive readiness that have been discussed in other chapters in this volume:

- *Physical health and well-being*: sufficient to be in school and have the fine motor, hand, and eye coordination to participate in activities, use scissors, hold a crayon, paint, paste, etc.
- *Social and emotional development*: the disposition and ability to get along with others, share, work in groups, be independent, and be able to be separated from home
- *Orientation and motivation for learning*: enthusiasm for learning some basic metacognitive skills, such as the ability to listen to and respond to instruction
- *Language development*: oral language skills and the background knowledge to associate words with things and the ability to communicate with others

- *Cognition and general knowledge*: for example, basic number and letter recognition; experience with words, objects, and natural environment; and the precursors to content knowledge

While physical health and well-being have not received much attention in other chapters, certainly physical adeptness seems a critical capacity for military readiness.

These then are the foundational skills that students need to be successful in kindergarten and to make good progress, and the establishment of these standards has encouraged important stakeholders to take note and to take action. Many parents and early childhood programs have responded by focusing on helping children develop the prerequisites they need for school success; some public agencies (e.g., state departments of education and school districts) have developed readiness tests for school entry. At the same time, however, there is general consensus, most recently observed by the National Research Council (Snow & Van Hemel, 2008), that the attention to school readiness has encouraged early childhood programs and assessment to overemphasize academic precursors and content skills at the expense of social and emotional development, motivation for learning, opportunities for self-expression, and more play-oriented approaches to learning and socio-emotional functioning.

Early childhood researchers and expert practitioners have called for a more balanced approach to program development and evaluation (Snow & Van Hemel, 2008). Recognizing that even in kindergarten, important assessments can drive and narrow practice, experts agree that effective early childhood assessments need to be comprehensive in addressing all important aspects of child development and that such assessments need to be integrated into the larger educational system that provides a strong infrastructure to support children's healthy development. I return to this point later in making recommendations for the military infrastructure needed to support the multiple dispositions and skills needed for cognitive readiness.

California's desired results system. For example, California has created its Desired Results System to guide the development, evaluation, and refinement of effective programs for early childhood and youth development; to identify exemplary programs and practices as well as focus programs and their attention to the assessment, screening, and support of students' learning progress (CDE, 2009). In contrast to prior evaluation models that emphasized program inputs, resources, and processes, *Desired Results* is intended to focus on programs and make them accountable for their results for children in four developmental categories and for those families in two categories:

- Children are personally and socially competent.
- Children are effective learners.
- Children show physical and motor competence.
- Children are safe and healthy.
- Families support their children's learning and development.
- Families achieve their goals.

Each of these outcome areas is defined by multiple indicators that encompass developmental themes. In turn, it is assessed through multiple measures, using standardized measurement tools that quantify and classify students' status in each area and classify their developmental levels. In addition to the tools and reporting mechanism, the system also provides for professional development to enable educators to understand and use the system. The system is available online and is used by state, regional, and local programs to evaluate quality and support progress toward successful results.

6.3 Readiness for College

Like readiness for kindergarten, the concept of readiness for college also is undergoing transformation. Traditional readiness indicators centered on what students needed for college admission: typically, such indicators as having a high school diploma and having taken and passed required college preparatory courses (e.g., 3–4 years of mathematics, English, foreign language, and science), grade point averages (GPAs), and scores on admissions tests such as the SAT and ACT.

However, as I discuss later, this definition has not served students well. Instead, advocates are promoting a new definition of what college readiness means and promoting its implications for K-12 education, i.e., readiness for college means not only possessing the entering requirements but more importantly having the knowledge, skills, and disposition to effectively transition to the college environment, to be ready to take credit-bearing courses, and to be successful in completing college.

College-ready capabilities. As with kindergarten readiness, there are many lists of what students need to be prepared for college in this sense. For example, Achieve, the Fordham Foundation, and a handful of state or university systems have laid out what they deem to be essential entering competencies (see, e.g., Achieve, 2008a, 2008b; 2009; ACT, 2005; CCC, CSU, and UC, 1982). David Conley (2007), a pioneer in promoting the alignment between K-12 and higher education and a prominent voice in the movement, has conducted research through surveys of college faculty, analysis of college course syllabi, and requirements to define the capabilities that students need to be college ready as:

- *Cognitive strategies* are intellectual behaviors and cognitive capabilities students need to engage with, understand, and master rigorous college-level content. These strategies include intellectual openness, inquisitiveness, analytic reasoning, argumentation, proof, interpretation, precision and accuracy, and problem solving.
- *Content knowledge* is the overarching academic knowledge and skills that students should develop in high school: writing, research, understanding of key principles within disciplines, and an ability to apply content knowledge in a variety of contexts.
- *Academic behaviors* are basic learning capabilities and include such capacities as metacognition, ability to reflect on and respond to one's misunderstanding,

and persistence, ability to apply a range of learning strategies to improve performance: These include, as well, a strong work ethic and ability to work collaboratively and in teams.
- *College contextual skills and awareness*. These are the knowledge and skills that enable students to function comfortably in a college context, such as the knowledge of norms, values, and conventions of college interactions; human relations skills to cope in new environments; as well as specific knowledge about requirements for college and how to apply.

While Conley explicitly calls out cognitive strategies and content knowledge as key categories, he embeds metacognition, motivation, and teamwork in the "academic behaviors" category and also includes additional dimensions of social competence within "college contextual skills and awareness." He observes that all of these capabilities and competencies should be purposively developed and practiced in high schools. For example, students can learn to intentionally apply cognitive strategies in their academic and other work, and with practice, these strategies can become a habitual way of working. As with the kindergarten readiness set of knowledge and skills (previous noted), there also are interesting similarities between the mix of general and specific capabilities that Conley identifies and those that have been associated with cognitive readiness in other chapters of this book.

6.4 Readiness for Challenging Work: US Secretary of State

Taking the readiness for a challenging career to an extreme, let us consider what may be required to be the US Secretary of State. In a newspaper feature (Miller, 2008), former United States' Secretaries of State provided advice to the incoming Secretary, Hillary Rodham Clinton, implicitly responding to the issue: what does it take to be ready to be Secretary of State? What does it take to be ready for a most challenging work assignment, where problems of complex, propelled by multiple factors that are difficult to assess and beyond one's control and situations and outcomes, are dynamic, uncertain, and difficult to predict? What does it mean to have the knowledge and skills to be successful in context where patterns are difficult to interpret, the connections between reaction and response tortuous, and motives, primary characters, and available options are complex and changing? Where are social relationships, partnerships, and teaming essential and ever changing? There even may be uncertainty in what success means—is it the best possible outcome, the one that optimized on the greatest variety of conflicting interests, that which suffices short- or long-term goals, etc.? Certainly the capability to achieve deals with such uncertainty and meet fuzzy outcomes is harder to define than those of kindergarten or college readiness, but former Secretaries' words of advice sound familiar themes as they offered their perspectives the principal roles of the position:

- George P. Shultz: Be a vigilant gardener. As Shultz notes, as a gardener, "You have to pull weeds when they're small and keep track of things. The same is true

in diplomacy." He goes on to emphasize the ways in which Secretaries of States need to build agendas, establish relationships, and be clear on what you like and don't—both with countries across the globe, the Congress, and the president. He also mentions the importance of recruiting and maintaining talented staff, "You may be a brilliant general, but if you don't have good troops, you're not going to get anything done."
- James Baker: Most important is "having a seamless relationship with the president" and that you have an understanding that the president will "protect your backside" and goes to note the importance of clear, not mixed signals and of being able to build national consensus on tough, contentious issues.
- Madeleine Albright: "The thing that is most difficult is setting priorities" and being attuned on rapidly changing situations and being able to respond to multiple demands simultaneously and to sift through and respond nearly instantaneously to complex, evolving information. She also underscores the importance of putting together a good team who will keep things on track.
- Colin Powell: "Leadership and management are as essential to the job of Secretary of State as are foreign policy formulation, world travel, and dealing with the crises that come your way." Powell also notes the importance of building and tending to relationships with major alliances and countries and the time he had to learn, e.g., international economic issues, relationship, products, and policy.

6.5 Readiness Starts with Clarity in Goals

As complex as the context, the themes raised in the Secretaries' advice echo those that were evident for kindergarten and college readiness and those that have been highlighted in other chapters of this book:

- Content knowledge and skills and ability to apply and transfer them in new situations
- Cognitive strategies, including problem solving, critical thinking, and analytic reasoning
- Social competence, including teamwork and leadership
- Communication
- Motivation/work ethic
- Ability to learn/metacognition

Surely the nature, complexity, and breadth of knowledge, skills, and underlying dispositions vary enormously from one context to the next, but the basic categories remain consistent: to be ready for college, students need to understand basic concepts and principles in academic disciplines, while a Secretary of State may need to understand the intricacies of historical context, global economies, political interests, competing values, etc., to address specific context, but both rely on content knowledge. Kindergarten social skills may focus on being able to share one's toys, while college students need to be able to work collaboratively, and a Secretary of State needs to engage in global partnerships, but collaboration, consensus building, and teamwork are at the heart of each. Communication, abilities to learn and benefit from experience, and abilities to solve problems—these are all consistent themes in what it means to be ready for success, whatever the specific context. Situational awareness and abilities to respond quickly to new crises are apparent in Secretary of State responsibilities but have simpler counterparts on kindergarten and high school-to-college transitions.

In education, it is commonly accepted that educational effectiveness and school improvement start with being clear on expected outcomes and that expected outcomes articulate standards for student learning, (i.e., what students need to know and be able to do to be successful in the future) whether that future be the next grade, college, or other postsecondary training; the world of work; or in any other transition. Standards for student learning then are the foundation on which both programs and assessments are developed. The programs, which may be school curriculum, instruction, special interventions, or other initiatives, are developed to enable students to reach the standards, and the assessments measure how well students are doing relative to the standards and/or what progress they are making. The measures, in turn, are used to monitor, support, and hold stakeholders accountable for progress.

Taking a leaf from education, then preparing for readiness, starts with articulating standards for readiness and developing programs and measures to address the standards. While a simple model to articulate—specify goals, measure progress, and take action to ameliorate any gaps—current educational practice falls short of the ideal. The match between the assessments and the standards is essential to the success of the model, but imperfectly realized.

6.6 Consequences of Mismatch: Readiness Expectations Versus Readiness Assessment

Given that there is some consensus on what classes of knowledge and skill are needed to help assure a smooth transition from early childhood to kindergarten, school, or the world of work, how well do our assessments match them? Let us consider what happens when they do not, specifically the case of data on the success of decisions about college admissions. As mentioned earlier, rather than taking into account the knowledge and skills needed for success, the criteria for college admissions tend to be high school completion, completion of and grades in academic coursework, and in many cases, scores on the SAT/ACT.

In terms of success, existing statistics paints a disappointing picture (see Conley, 2008):

- Surveys of faculty, whether they be research institutions, state, or community colleges, routinely report students unprepared to meet their expectations for college coursework.
- More than 40 % of entering college students need remedial coursework in English and/or mathematics.

- Less than 20 % (17 %) of students enrolled in remedial reading upon college entry complete college.
- Only 60 % poor and minority students graduate high school. Of these students:
  - Only 1/3 enter college.
  - Less than half of those entering (1/7) complete college.
- The United States ranks fifteenth of 29 countries internationally in the proportion of students who complete college.

These data provide important feedback for the K-12 system. They also demonstrate the utility of multiple indicators for measuring readiness.

Moreover, when we look at the validity of the indicators used for college admissions, the power is quite weak. For example, the results from a validity study conducted by the University of California (UC) document the relationship between UC admissions criteria and first year college grades (Geiser & Studley, 2003). The study tracked 80,000 students admitted as freshmen to a UC campus from 1996 to 1999 and clearly suffers from range restriction. That is, the UC system is highly selected and intended to serve the top 12.5 % of California's high school graduates. Average high school GPA for the study sample was 3.5 and combined verbal and math SAT scores on average were nearly 1,200 (578 SAT verbal; 611 SAT math). While range restrictions attenuate the study's observed relationships, results show that SAT scores accounted for only 12.8 % of the variation in freshman GPA. The SAT II, tests that address more specific areas of content achievement, explained 15.3 % of the variation in freshman GPA, and high school GPA explained 14.5 % of the variation. The combined indicators accounted for little more than 21 % of the variation in freshman GPA, and interestingly, SAT I scores improved the prediction rate only by a negligible 0.1 % (from 21.0 to 21.1 %), making its value added a bit questionable. Furthermore, SAT I scores proved to be more susceptible to the influence of the socioeconomic status of an applicant than either the SAT II scores or high school GPA. That is, there was a higher relationship between students' socioeconomic status and their SAT I scores than between their such status and either performance on the SAT II or high school GPA stated in terms of correlation coefficients rather than percent of variance explained; the correlations between the individual indicators (SAT, SATII, high school GPA) and freshman GPA ranged from 0.36 to 0.39.

In short, there appears to be a problem in the alignment of the educational system and between the education and assessment systems. Indicators of college preparedness relative to the need for remedial course work and completion rates, especially for students of low socioeconomic status and students of color, suggest that precollegiate education, K-12, does not well support the development of knowledge and skills that students need for college success, and neither are assessments for college entry well aligned with what is needed for success. Clearly, graduating high school with a high school diploma does not translate into being ready for college expectations. In California, approximately 60 % of new freshmen in the California State University system are judged deficient in English, mathematics, or both and thus need to be assigned to remedial, noncredit-bearing English and/or mathematics courses (Chronicle of Higher Education, 2010), and such figures are even more severe for poor and minority students, e.g., Latino and African American. Nor do current high school proficiency tests translate into such preparation. For example, a number of studies have documented the thin relationship between expectations evident in high school exit exams and those of entry-level college courses and/or college performance (Brown & Conley, 2007; D'Agostino & Bonner, 2009). Similarly, student failure rates in Algebra I stand in stark contrast to college expectations for successful completion of Algebra II and Geometry.

Even for those who are successful in high school coursework, passing the so-called college preparatory requirements, such as the A-G course requirements for the University of California, does not add up to readiness for the rigor of being able to independently deal with the demands of college. And, in fact, we know that passing courses with the same name can mean something vastly different in terms of rigor, knowledge, and skills, depending on what school you go to (Achieve, 2008a, 2008b, ACT, 2005; U.S. Department of Education, 2007). Neither does high school generally prepare students for the social interactions and group work that are a core element of freshman coursework, according to national college surveys (Conley).

6.7 Moving to a More Aligned System

In K-12 we are recognizing the need to move from the current chaos to a coherent system of development and measurement, a system that is horizontally and vertically aligned with expectations for student performance and particularly that is aligned with the knowledge and skills students need to be well prepared for college and the world of work.

Fewer, clearer, higher standards. Step one of the planned move is agreement on core standards that define what students need to know and be able to do to be prepared for college. This new generation of standards is intended to correct the shortcoming of current standards—standards that were too vague and too many to provide a strong foundation for education and schooling. In its place is a call for "Fewer, Clearer, Higher" (FCH) standards, meaning in general that standards should be (see Common Core State Standards Initiative, 2010; Herman, 2009):

- *Fewer*: Represent a powerful and coherent set of *essential competencies* that all students can be expected to develop over the course of their K-12 education to be college and/or work ready, reasonable, yet still cognitively demanding in scope such that all students can be expected to acquire them to graduate high school.
- *Clearer:* Be sufficiently clear to guide the development of assessments to support accountability and improvement for students, educators, administrators, and the system as a whole; be sufficiently clear to guide the design and provision of rigorous coursework to enable students to achieve such competencies; and explicitly define expected levels of content and cognitive demand.
- *Higher*: Represent the knowledge, skills, and competencies that students need to be prepared for success in college and the workplace; incorporate deep conceptual understanding and high levels of cognitive demand, including abilities to apply

and transfer knowledge, reason, conduct inquiry, and communicate; and be benchmarked to the international standards.
- *Defensible*: Meet criteria for fairness, credibility, and accuracy.

Common Core State Standards have been developed nationally for English language arts and mathematics. The first set of documents defined high school graduation expectations that would prepare students for success in college and work. Following vetting of that specification, grade-level expectations were articulated for each grade, K-12, to progressively build to the knowledge and skills required for high school graduation and subsequent success. The standards development assumes that the same knowledge and skills are needed regardless of whether students plan to enter and complete college or to enter the workforce more directly, in that livable wage jobs require postsecondary technical training and that success with such requires similar mathematics and English language arts capability to college readiness.

In any event, Common Core State Standards are intended to provide a strong, sound foundation for systems that support accountability and the improvement of learning. A prominent national commission described such systems as needing to be comprehensive, addressing all the important learning goals; coherent across levels of the educational system and across grade levels; and continuous, providing ongoing information on how students were doing (Pellegrino, Chudowsky, & Glaser, 2001).

System coherence is a hallmark, in contrast to current practice where ongoing classroom assessment and large-scale assessment tend to be disjointed and, as noted, K-12 and college expectations are frequently out of sync (Conley, 2007; Herman, 2010). Instead, standards and assessment must be *horizontally coherent*, meaning a system where curriculum, instruction, and assessment are aligned, so that assessment results can be used to judge progress and learning effectiveness as well as to guide instruction. *Vertical coherence* brings into line all levels of the educational system—classroom, school, district, and state, so that all share the same understanding of the goals for student learning and how it is expected to develop over time so that all system resources—funding, leadership, mentoring, professional development, special programs, technical assistance, curriculum, instruction, and assessment, support a unified effort to building students' capacity.

Finally, *developmentally coherent* means that the system reflects a continuous view of learning and how it is expected to develop over time, for example, from the beginning to the end of the year and across years, grades, and levels of schooling. *Developmental coherence* means that subject area standards for student learning—mathematics, for example—progressively build from one grade level to the next to articulate a logical progression of the knowledge and skills students need to develop in each grade to be able to graduate high school and be prepared for success in college and work. In a developmentally coherent system, expectations for each grade level progressively build from the prior to the next level to directly map to ultimate goals.

A coherent, standards-based system communicates a consistent set of targets for all stakeholders in the system—e.g., students, teachers, administrators, and policymakers; aligns standards, curriculum, and assessment to assure that students have the opportunity to learn what is expected, that which is required for future success;

and systematically builds and supports students' continuous progress toward immediate and ultimate goals. Assessment serves as a critical lynchpin in such a system: whether large-scale accountability, summative, or classroom assessments, assessment communicates to students what is important to know and be able to do and what is valued knowledge; it also provides evidence through which to judge the status and/or progress of learning and on which to base subsequent action, for example, to strengthen programs, identify students who need help, and to suggest next steps to facilitate progress (Herman, 2008). While large-scale accountability assessments have traditionally had the lion's share of attention in research and policy, research suggests that the most powerful use of assessment to improve learning comes not from large-scale state assessments, but from the ongoing formative use of assessment guide immediate teaching and learning (Black & Wiliam, 1998, 2004; OECD, 2005; Phelan, Choi, Vendlinski, Baker, & Herman, 2009). That is, assessment during the course of instruction or training, when it diagnoses students' learning needs and is used to inform the next steps for teaching and learning, shows strong, positive effects on learning, particularly for low ability students.

6.8 Systems for Developing Cognitive Readiness

These same issues apply to attempts to develop cognitive readiness for the military. Based on current work in the K-12 sector, reasonable solutions may lie in:

Clarifying and clearly communicating expectations for preparation and readiness. Research on standards and assessments in K-12 education demonstrates that standards and assessments make a difference; committed stakeholders listen to the signal they send and take action. Consider, for example, the unveiling of school readiness indicators and the explosion of games of "I spy" to encourage young children to recognize and use words and develop language as well as massive movement in preschool and afterschool programs to develop academic skills.

Systematically integrating opportunities for readiness development. Being clear on readiness expectations facilitates attention to their development. K-12 research clearly shows that educators respond to high visibility standards and assessment by focusing their curriculum and teaching on what is assessed and that curriculum developers, publishers, and other service providers respond by adapting their existing materials to address new goals and/or by creating new materials and services that address them (Herman, 2007).

Here is an area where the military could build from historic shortcomings in K-12 practice, where the process of articulating new goals and standards has tended to be additive rather than integrative. That is, standards are developed separately by subject—reading, mathematics, and science—and each set is the subject of intense and separate materials development, professional development, and implementation efforts. The result can be more standards, instructional materials, and assessments than there is available time to implement them and an overwhelming number of things for teachers—particularly elementary school teachers who teach multiple subjects—to instructionally implement, manage, and/or monitor. Alternatively, some recent initiatives have shown the value of a more integrative approach. For example, the Lawrence Hall of Science's *Seeds of Science/Roots of Reading Program* integrates the development of literacy and science knowledge within a common curriculum. Students use science as a context for learning to read and write and develop inquiry skills across both subjects. Not only does the program get science back into the curriculum (strong accountability requirements for schools' reading and mathematics performance have reduced attention to science), but program results show clear advantage for both literacy and science learning (Cervetti, Bravo, Hiebert, Pearson, & Jaynes, 2009).

Similarly, for military applications aimed at developing cognitive readiness, it does not seem wise to treat those goals separate from the content and mission skills in which they are naturally embedded. Cognitive readiness needs to be developed in the context of developing specific mission knowledge and skills. While cognitive readiness implies a set of common intellectual processes that individuals use to deal with unexpected conditions, the concept of cognitive readiness should be overlaid on the development of capability to deal with expected conditions in particular domains and be purposively designed to provide opportunities for near and far transfer.

A parallel in education is building students' problem-solving capacity, which research shows cannot be separated from the development or assessment of content understanding—i.e., be able to address novel, complex problems requires prior knowledge and deep understanding of the concepts, principles, procedures, etc., that need to be brought to bear and combined to solve the problem (see Mayer, this volume). Rather, content knowledge and problem-solving transfer need to be developed in tandem, and training needs to incorporate research-based principles for fostering transfer, such as providing practice over a range of contexts, use of analogical reasoning, developing and invoking explicit problem-solving strategies, e.g., recognizing and representing the problem, and identifying possible solution strategies (Chi, Glaser, & Farr, 1988; Holyoak, 2005; Koedinger & Corbett, 2005).

Define and test developmental trajectories for cognitive readiness. While research on problem solving suggests that some aspects of cognitive readiness cannot exist independent of content knowledge or of the domain for which it is to be exercised, it seems reasonable that it may be possible to develop domain-independent models of these specific skills that then can be instantiated within specific domains and learned repeatedly and can transfer to new domains of learning (Baker, 2007a, 2007b)—e.g., communicating clearly requires knowledge of what one is communicating about and knowledge of the conventions of communication in that domain, but learning to well communicate in one domain may generalize to learning to communicate in a second domain, given appropriate content knowledge in the second domain, and/or make subsequent learning faster and more efficient, similarly with problem solving, metacognition, and other components of cognitive readiness. Given such domain-independent models, it would be valuable to hypothesize and verify the trajectories through which such skills develop and then consistently embed them within and across training contexts. Systematically attending to problem solving within a training domain could both strengthen capability within that domain and support transfer of problem solving across training domains, i.e., problem-solving capability developed in one domain could be transferred to benefit both content and problem-solving learning in subsequent domains.

Attention to transfer. Transfer is inherent in any definition of cognitive readiness: that is, officers and/or enlisted men must be able to draw on their knowledge and skills to meet novel circumstances and respond to unpredictable challenges. Theory and research in educational psychology suggest that transfer is enhanced when there is explicit attention to it in instructional sequences; training, for example, should provide students the opportunity to apply their knowledge across a wide range of contexts and to see the connections between new and prior problem contexts (Chi et al., 1988; Holyoak, Gentner, & Kokinov, 2001; Koedinger & Corbett, 2005). There is an inherent tension here between the breadth and depth of knowledge and skills that can be developed during finite period of instruction, for example, a broad survey of a topic (e.g., western civilization) may give scant opportunity for students to apply and transfer their knowledge in specific areas (e.g., capitalism). The movement in K-12 to FCH mentioned previously, at least in part, is an attempt to remedy this tension by being clear on the concepts and principles that are most important for students to be able to apply *and* that can be feasibly developed within the time available for coursework. The intent is both to make learning expectations clear and to enable teachers and schools to concentrate on deeper development of fundamental knowledge and skills that will prepare students for future success. In contrast, today's panoply of standards and learning objectives has produced a curriculum that is a mile wide and an inch deep (Schmidt, Houang, & Cogan, 2002), which is counterproductive to the deep learning needed for transfer.

The lack of transfer also is evident in comparisons between students' scores on high visibility, state accountability tests, and those on other measures of the same subject at the same grade levels. That is, if meaningful learning has occurred, one would expect scores on one test to generalize to those on another, similar test of the domain. Instead, in the K-12 world, Koretz (2008) shows that students' scores on Kentucky's state assessment showed steep improvement over the 1990s in reading and mathematics, yet the National Assessment of Educational Progress showed little, if any, gains in learning and raises the specter of score inflation.

Those scores do not generally point to problems in both instruction and assessment, neither is sufficiently attentive to transfer. Research shows that teachers teach to the test and teach like the test – i.e., engage students in exercises that model test formats and are limited to test content—not to students' ability to understand and apply what they learn across multiple contexts (see, e.g., Hamilton, Stecher, Russell, Marsh, and Miles, 2008; Herman, 2004). More explicit attention to transfer in both the development and assessment to training could help to alleviate these shortcomings.

Coherent systems of readiness assessment. The military should consider the development of a coherent system for the development and assessment of cognitive readiness grounded in an explicit definition of the cognitive readiness construct and how it may develop in the context of mission-oriented knowledge and skills. Such a system would be based on multiple indicators of relevant domains and yield the formative and summative information needed to monitor, improve, and certify performance as well as to support continuous development and the alignment of current training to subsequent expectations. California's Early Assessment Program (CSU, 2009), a joint assessment initiative of the State Board of Education, the California Department of Education, and the California State University (CSU) system provide one example of such a multipurpose system. Coming together to examine the alignment between K-12 standards and college expectations, the partners augmented the California Standards Tests (CST) of English and mathematics given to all 11th graders as part of the state accountability system to create a supplement that, combined with selected CST items, provides a reliable estimate of student performance relative to CSU's readiness expectations. The intent is to give students early feedback on their readiness to take credit-bearing—rather than remedial—courses upon college entry and to provide a strong signal to teachers and schools about what they need to do to better prepare students for college success. Participation in the supplementary testing is voluntary for students; those who participate get feedback on whether they have met CSU expectations, and thus are exempt from further placement exams at college entry; those who score "not ready" are informed that they need additional preparation for college-level work. Such preparation presumably is to occur during students' senior year in high school, also serving the purpose of combating the "senior slump," that is, the tendency for students to tune out of school after their college applications and grade submissions. Students whose scores show the need for additional preparation can take a variety of diagnostic tests to determine individual strengths and weaknesses to support their subsequent development and can access resources and modules to support their development on the state website. Grade 12 coursework also was redesigned and standard sequences, professional development, and instructional resources provided to support teacher implementation. In addition, CSU expectations, data on the strengths and weaknesses of student performance, and strategies for dealing with them also were incorporated into CSU teacher preparation programs. Expectations for student preparation thus provide a touchstone for more closely aligning K-12 education and college entry expectations for constructing a complementary system of assessments to provide accountability; formative and diagnostic assessment data for students and high school educators; placement data for CSU; and feedback for strengthening the pre- and in-service preparation of high school teachers to develop their students' college readiness skills in English language arts and mathematics. Not only a coordinated system of *assessments*, the system also includes model courses and instructional packages for diagnosing and responding to students' strengths and weakness relative to college preparation.

The senior year of high school may not provide sufficient time to remediate the shortcomings of the prior 11 years, and in fact, as mentioned previously, the K-12 community is currently in the process of reconfiguring its expectations of K-12 so schools will help students, grade by grade, to develop systematically the knowledge and skills students need to be prepared for college success. Nonetheless, the Early

Assessment Program demonstrates some of the many linkages through which a comprehensive assessment system can support learning and the many actors and venues through which change must occur.

Similarly, if the military wants to build the cognitive readiness of its forces, it would do well to make its expectations clear and incorporate the development and assessment of specific, relevant requisite knowledge, skills, and dispositions into all training programs. By regularly assessing individual status with regard to cognitive readiness, the military could signal the importance of these capacities to trainers and trainees alike and support accountability and provide important feedback on whether these capacities are being sufficiently developed. The feedback could be used to improve readiness within and across required training as well as to enable individuals to take responsibility for alleviating observed shortcomings. Diagnostic batteries and effective instructional resources linked to desired capacities would help to complete the system.

6.9 Summary and Conclusions

The K-12 public education sector offers a number of lessons that may transfer to the military's mandate to develop the cognitive readiness of its officers and enlisted personnel. A first issue is definitional: what does "cognitive readiness" mean and how does it develop? Conceptions of kindergarten readiness, college readiness, and the capacities needed for complex careers suggest many parallels to the conceptions of cognitive readiness articulated by other chapters in this volume. Being clear on expected capacities is a clear first step to assuring that such cognitive readiness capacities are systematically developed throughout education and training experiences.

Coursework must provide trainees adequate opportunities to develop the knowledge, skills, predispositions, and transfer capabilities that define cognitive readiness. Just as with problem solving in the K-12 sector, it is unlikely that cognitive readiness can be developed independent of domain knowledge, that is, independent of mission-oriented content. Rather, opportunities to develop and apply cognitive readiness skills and dispositions must be embedded systematically within and across courses addressing the development of mission-oriented domain knowledge and skills. Coursework instruction and assessment should address trainees' ability to apply and transfer their knowledge to new contexts and dynamic circumstances.

Experience in K-12 also demonstrates important advances in creating aligned systems for developing and assessing readiness knowledge and skills. Such systems take a developmental perspective of how competencies develop and align instruction and assessments to significant benchmarks along the way. A variety of coordinated assessments support student development: accountability or summative assessments which are targeted on expected course outcomes, formative assessments that address the subsidiary knowledge and skills which help to fuel the outcomes, and diagnostic assessments which may identify specific strengths and weaknesses with regard to specific knowledge and skills: they mutually focus teachers and students on a coherent set of goals and sequence of learning and provide suitable data for the range of decisions that support quality in teaching and learning. That the assessment system is closely coordinated with expected instructional sequences provides ongoing information to guide instructional decision making, diagnose and respond to individual needs, and to support reflection and improvement of teaching.

Assessment and the improvement of cognitive readiness: yes, the military can learn valuable lessons from K-12, but K-12 should also benefit from the military efforts to reach its goals.

Acknowledgements The work reported herein was partially supported by grants from the Office of Naval Research, Award Number N000140810126, and the Gates Foundation, Award Number 59559. The findings and opinions expressed here do not necessarily reflect the positions or policies of the Office of Naval Research or the Gates Foundation.

References

Achieve, Inc. (2008a). *Out of many, one: Toward rigorous common core standards from the group up*. Washington, DC: Author.

Achieve, Inc. (2008b). *Closing the expectations gap 2008*. Washington, DC: Author.

Achieve, Inc. (2009). *Ready or not: Creating a high school diploma that counts*. Retrieved December 18, 2009, from http://www.achieve.org/node/175

ACT, Inc. (2005). *Crisis at the core: Preparing all students for college and work*. Iowa City, IA: Author.

Baker, E. L. (2007a). The end(s) of testing (2007 AERA presidential address). *Educational Researcher, 36*(6), 309–317.

Baker, E. L. (2007b). Model-based assessments to support learning and accountability: The evolution of CRESST's research on multiple-purpose measures. *Educational Assessment (Special Issue), 12*(3&4), 179–194.

Biemiller, A. (2006). Vocabulary development and instruction: A prerequisite for school learning. In S. Newman & D. Dickerson (Eds.), *Handbook of early literacy research*. New York: Guillford Press.

Black, P., & Wiliam, D. (1998). Inside the black box: Raising standards through classroom assessment. *Phi Delta Kappan, 80*(2), 139–148.

Black, P., & Wiliam, D. (2004). The formative purpose: Assessment must first promote learning. In M. Wilson (Ed.), *Towards coherence between classroom assessment and accountability* (pp. 20–50). Chicago, IL: University of Chicago Press.

Brown, R., & Conley, D. (2007). Comparing high school assessments to standards for success in entry-level university courses. *Educational Assessment, 12*(2), 137–160.

California Community Colleges, California State University, University of California. (1982). *Statement on competencies in mathematics expected of all entering freshman*. Retrieved December 18, 2009, from http://www.universityofcalifornia.edu/senate/reports/mathcomp.html

California Department of Education (CDE). (2009). *Introduction to desired results*. Retrieved October 13, 2009, from http://www.cde.ca.gov/sp/cd/ci/desiredresults.asp

California State University (CSU). (2009). *Early assessment program*. Retrieved December 19, 2009, from http://www.calstate.edu/EAP/12/19/09

Cervetti, G. N., Bravo, M. A., Hiebert, E. H., Pearson, P. D., & Jaynes, C. (2009). Text genre and science content: Ease of reading, comprehension, and reader preference. *Reading Psychology, 30*(6), 487–511.

Chi, M. T. H., Glaser, R., & Farr, M. J. (Eds.). (1988). *The nature of expertise*. Hillsdale, NJ: Erlbaum.

Chronicle of Higher Education. (2010, March 18). Cal State to require remedial courses before freshman year. *Chronicle of Higher Education*. Retrieved January 14, 2011, from http://chronicle.com/blogs/ticker/cal-state-to-require-remedial-courses-before-freshman-year/21915

Common Core State Standards Initiative. (2010). *Common Core State Standards Initiative: The standards*. Retrieved January 14, 2011, from http://www.corestandards.org/the-standards

Conley, D. T. (2007). *Redefining college readiness*. Report prepared for the Gates Foundation. Eugene, OR: EPIC.

D'Agostino, J., & Bonner, S. (2009). High school exit exam scores and college performance. *Educational Assessment, 14*(1), 25–37.

Farkas, G., & Beron, K. (2004). The detailed age trajectory of oral vocabulary knowledge: Differences by class and race. *Social Science Research, 33*, 464–497.

Fletcher, J. D. (2004). *Cognitive readiness: Preparing for the unexpected*. Alexandria, VA: Institute for Defense Analysis.

Geiser, S., & Studley, R. (2003). UC and the SAT: Predictive validity and differential impact of the SAT I and SAT II at the University of California. *Educational Assessment, 8*(1), 1–26.

Hamilton, L. S., Stecher, B. M., Russell, J. L., Marsh, J. A., & Miles, J. (2008). Accountability and teaching practices: School-level actions and teacher responses. *Research in the Sociology of Education, 16*, 31–66.

Hart, B., & Risley, T. (1995). *Meaningful differences in the everyday experience of young American children*. Baltimore, MD: Brookes.

Herman, J. L. (2004). The effects of testing on instruction. In S. Fuhrman & R. Elmore (Eds.), *Redesigning accountability*. New York: Teachers College Press.

Herman, J. L. (2007). *Accountability and assessment: Is public interest in K-12 education being served?* (CRESST Report 728). Los Angeles, CA: University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Herman, J. L. (2008). Accountability and assessment in the service of learning: Is public interest in K-12 education being served? In L. Shepard & K. Ryan (Eds.), *The future of test based accountability*. New York: Taylor and Francis.

Herman, J. L. (2009). *Moving to the next generation of standards for science: Building on recent practices* (CRESST Report 762). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Herman, J. L. (2010). *Coherence: Key to next generation assessment success (AACC Report)*. Los Angeles, CA: University of California.

Holyoak, K. J. (2005). Analogy. In K. J. Holyoak & R. G. Morrison (Eds.), *The Cambridge handbook of thinking and reasoning* (pp. 117–142). Cambridge, England: Cambridge University Press.

Holyoak, K. J., Gentner, D., & Kokinov, B. N. (2001). Introduction: The place of analogy in cognition. In D. Gentner, K. J. Holyoak, & B. N. Kokinov (Eds.), *The analogical mind: Perspectives from cognitive science* (pp. 1–19). Cambridge, MA: MIT Press.

Koedinger, K. R., & Corbett, A. (2005). Cognitive tutors: Technology bringing learning science to the classroom. In K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (pp. 61–78). Cambridge, England: Cambridge University Press.

Koretz, D. (2008). *Measuring up: What educational testing really tells us*. Cambridge, MA: Harvard University Press.

Lee, V., & Burkham, D. (2002). *Inequality at the starting gate: Social background differences in achievement as children begin school*. Boston: Economic Policy Institute.

Miller, M. (2008, December 7). Speaking from experience. *Los Angeles Times*. Retrieved from http://articles.latimes.com/2008/dec/07/news/OE-MILLER7

National Commission on Excellence in Education. (1983). *A nation at risk: The imperative for educational reform*. Washington, DC: U. S. Government Printing Office.

OECD. (2005). *Formative assessment: Improving learning in secondary classrooms*. Paris: Author.

Pellegrino, J. W., Chudowsky, N., & Glaser, R. (2001). *Knowing what students know: The science and design of educational assessment*. Washington, DC: National Academy Press.

Phelan, J., Choi, K., Vendlinski, T., Baker, E. L., & Herman, J. L. (2009). *The effects of POWERSOURCE intervention on student understanding of basic mathematical principles* (CRESST Tech. Rep. No. 763). Los Angeles: University of California, National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

Schmidt, W. H., Houang, R. T., & Cogan, L. S. (2002). A coherent curriculum: The case of mathematics. *American Educator, 26*(2), 10–26.

Snow, C., & Van Hemel, S. (Eds.). (2008). *Early childhood assessment: Why, what, and how*. Washington, DC: NRC.

U.S. Department of Education. (2007). *America's high school graduates: Results from the 2005 NAEP high school transcript study*. Washington, DC: National Center for Education Statistics, NCES.

Chapter 7
Cognitive Readiness for Solving Equations

Gregory K.W.K. Chung and Girlie C. Delacruz

Algebra plays a significant role in a student's academic pathway, from being a graduation requirement to acting as a gatekeeper to more advanced coursework required of science, technical, engineering, and mathematics majors and careers. Yet, the magnitude of students' poor math preparation is staggering at every level of schooling. For example, in the Los Angeles Unified School District, the second largest school district in the United States, the pass rate for the California High School Exit Exam (CAHSEE) in math was 47 % (CDE, 2011) and mirrors the weak performance of eighth graders on the 2009 National Assessment of Educational Progress (NCES, 2010). Poor math preparation has resulted in large numbers of students failing the California State University math entrance exam (16,900 or 35 % of the fall 2010 first-time freshmen; CSU, 2011). The consequence of poor math preparation is of national importance, as failure to maintain a pipeline of prepared students is diminishing the nation's competitive technology infrastructure and lead in the global economy (NAE & IOM, 2006; NAS, 2005; NSB, 2010). In our own studies with middle school students, we have found about 50 % of eighth-grade students did not recognize that $12 \times (1/12)$ equals 1, and about 34 % of these students could not provide a solution to the equation $3x + 1 = 13$ (Chung et al., 2007). Success in algebra is predicated on students developing foundational math concepts and skills. The National Mathematics Advisory Panel (NMAP, 2008) identified *fluency with whole numbers* as a critical skill underlying algebra. Fluency refers to the ease with which learners can manipulate whole numbers quickly and with automaticity.

O'Neil's cognitive readiness learning model (O'Neil, Lang, Perez, Escalante, & Fox, this volume) comprises a domain-independent set of knowledge, skills, and attributes. The knowledge component is conceptualized as the prerequisite

G.K.W.K. Chung (✉) • G.C. Delacruz
National Center for Research on Evaluation, Standards, and Student Testing (CRESST)/
University of California, Los Angeles (UCLA), 10945 Le Conte Avenue, Ste. 1355,
Box 957150, Los Angeles, CA 90095-7150, USA
e-mail: greg@ucla.edu; gdelacruz@cse.ucla.edu

H.F. O'Neil et al. (eds.), *Teaching and Measuring Cognitive Readiness*,
DOI 10.1007/978-1-4614-7579-8_7,

knowledge and the domain-specific knowledge needed to develop cognitive readiness in the domain. The skills component includes the adaptability, adaptive problem solving, communication, decision making, and situation awareness. The attributes/competencies component includes adaptive expertise, creative thinking, metacognition, and teamwork.

This chapter examines the prerequisites for cognitive readiness that underlie solving equations, a fundamental topic in algebra. The prerequisite for solving equations exists under the knowledge component in O'Neil et al.'s cognitive readiness learning model. The prerequisites for solving equations require fluency with the basic operations (addition, subtraction, multiplication, division), use of properties (e.g., commutative, associative, and distributive properties), and the knowledge of how to apply these operations to problem solving. In the remainder of this chapter we first illustrate the complexity of solving equations by enumerating the set of mathematical operations that are needed to solve an equation. We then report on an assessment technique we have been investigating to measure cognitive readiness for solving equations.

7.1 What Knowledge Is Required to Solve an Equation?

In the context of school algebra, being cognitively ready to solve multistep equations suggests a learner possesses the knowledge of the properties of operations on numbers (often referred to as the properties of arithmetic or algebra), skill in applying a particular operation constrained by its properties, reasoning during the simplification (or transformation) process, and monitoring of the solution. Students need to be fluent with a variety of fundamental concepts such as negative numbers and the use of letters as an unknown (Carraher & Schliemann, 2007; Chazan & Yerushalmy, 2003; Vlassis, 2002).

A cognitive task analysis conducted by Chung et al. (2007) identified over 50 concepts related to solving equations. Students need to be facile with a substantial body of knowledge, including the meaning of equality; unit/1; the properties of algebra; rational numbers and integers; theorems and conventions (e.g., $-1 \times -1 = 1$; order of operations); the operations; and factorization.

Table 7.1 lists the general types of knowledge identified from the cognitive task analysis to underlie solving equations and Table 7.2 lists the properties of

Table 7.1 Sample of knowledge related to solving multistep equations at the middle school level

Numbers	Rational numbers (e.g., fractions, percentage, mixed, decimal, ratio)
Symbols	Operators (e.g., +, −, •, ×, /, ÷, fraction bar), grouping (e.g., parentheses, brackets, braces, absolute value, repeating decimals, square root), equality (=), sign (+, −), variables (e.g., *a*, *b*, *x*)
Operations	Addition, subtraction, multiplication, division
Conventions	Order of operations, grouping, and simplification (e.g., $(x) = x$)

Table 7.2 Properties of operations on real numbers

Property	Form
Identity properties of addition and multiplication	$a+0=a \quad a\bullet 1=a$
Associative property of addition and multiplication	$a+(b+c)=(a+b)+c \quad a\bullet(b\bullet c)=(a\bullet b)\bullet c$
Commutative property of addition and multiplication	$a+b=b+a \quad a\bullet b=b\bullet a$
Additive and multiplicative inverse properties	$a+(-a)=0 \quad a\bullet(1/a)=1, a\neq 0$
Distributive property	$a\bullet(b+c)=a\bullet b+a\bullet c$
Addition property of equality	if $a=b$, then $a+c=b+c$
Multiplication property of equality	if $a=b$, then $a\bullet c=b\bullet c$
Reflexive property of equality	$a=a$
Symmetric property of equality	if $a=b$, then $b=a$
Transitive property of equality	if $a=b$ and $b=c$, then $a=c$
Theorems	$(-1)\bullet(-1)=+1 \quad (-1)\bullet(+1)=-1$

operations on real numbers. Thus, depending on the knowledge of and fluency with these concepts, solving equations can range from being automated and error-free to being deliberate and error-prone.

7.2 The Cognitive Demands of Solving Equations

To illustrate the cognitive complexity of solving an equation, we conducted a cognitive task analysis on the equation $7x-(3x-2)=38$. This equation is typical of what students would receive in an algebra class. While the solution is straightforward, we have observed a low percentage of students in algebra—particularly those struggling in math—unable to solve this equation, with less than 15 % of our sample successfully solving this equation (Chung et al., 2007). What is puzzling about this low performance is that the equation is simple and straightforward, and only requires knowledge of math concepts drawn from Tables 7.1 and 7.2 that presumably have been covered in the elementary math curriculum.

Our analysis focused on uncovering the underlying mathematical operations required to solve $7x-(3x-2)=38$. Our analysis technique was to solve the equation step-by-step, such that each transition from one step to the next step used only *one* mathematical operation. Table 7.3 shows one solution path for the equation $7x-(3x-2)=38$. The solution steps use a single operation from Table 7.1 or 7.2. Overall, the steps in Table 7.3, when categorized in terms of the categories listed in Table 7.1 or 7.2, have nine properties of operations, one theorem, eight arithmetic operations, and two simplifications. Of particular interest are two critical transitions: Step 11 → 12 and Step 15 → 16. These transitions are critical because they are the only ways to simplify the equation. Less obvious examples include Step 13 → 14 and Step 17 → 18, of which the underlying mathematical reason (identity properties) for the 0 or 1 "vanishing" from the expression is often a mystery to students.

Table 7.3 Solution steps for a multistep equation

Step	Equation	Type of operation used in transforming equations shown in the "Equation" column[a]
Given	$7x-(3x-2)=38$	Definition of subtraction: $a - b = a + (-b) \rightarrow a = 7x$; $b = (3x - 2)$
1	$7x+(-(3x-2))=38$	Theorem: $(-1) \bullet (+1) = -1 \rightarrow -(3x - 2) = (-1)(3x - 2)$
2	$7x+(-1)(3x-2)=38$	Definition of subtraction: $a - b = a + (-b) \rightarrow a = 3x$; $b = 2$
3	$7x+(-1)(3x+(-2))=38$	Distribution over addition: $a \bullet (b + c) = a \bullet b + a \bullet c \rightarrow a = -1$; $b = 3x$; $c = -2$
4	$7x+(-1)3x+(-1)(-2)=38$	Multiplication: $(-1) \bullet 3 = -3$
5	$7x+(-3)x+(-1)(-2)=38$	Multiplication: $(-1) \bullet -2 = 2$
6	$7x+(-3)x+(2)=38$	Simplify: $(2) = 2$
7	$7x+(-3)x+2=38$	Distribution (factor x): $a \bullet (b + c) = a \bullet b + a \bullet c \rightarrow a = x$; $b = 7$; $c = -3$
8	$x(7+(-3))+2=38$	Addition: $7 + -3$
9	$x(4)+2=38$	Simplify: $(4) = 4$
10	$x4+2=38$	Commutative property of multiplication: $a \bullet b = b \bullet a \rightarrow a = x$; $b = 4$
11	$4x+2=38$	Addition property of equality: $a + c = b + c \rightarrow a = 4x + 2$; $b = 38$; $c = -2$
12	$4x+2+(-2)=38+(-2)$	Additive inverse: $a + -a = 0 \rightarrow a = 2$
13	$4x+0=38+(-2)$	Identity property of addition: $a + 0 = a \rightarrow a = 4x$
14	$4x=38+(-2)$	Addition: $38 + -2$
15	$4x=36$	Multiplication property of equality: if $a = b$, then $a \bullet c = b \bullet c \rightarrow a = 4$
16	$(\frac{1}{4})4x=(\frac{1}{4})36$	Multiplicative inverse: $a \bullet 1/a = 1$, $a \neq 0 \rightarrow a = 4$
17	$(1)x=(\frac{1}{4})36$	Multiplicative identity: $1 \bullet a = a \rightarrow 1x = x$
18	$x=(\frac{1}{4})36$	Multiplication: $(1/4) \bullet 36$
19	$x=\frac{36}{4}$	Division: $36 / 4$
20	$x=9$	

[a]Each equation in the "Equation" column represents a step in the solution path. When an equation is operated on as described in the "Type of operation ..." column, the resulting transformed equation appears immediately under the original equation as the "next step." For example, given step 5, multiplying (–1) and (–2) in the step 5 equation results in the step 6 equation.

One advantage of deriving solution steps using only one operation is that it reveals all the underlying knowledge that is often chunked when one actually carries out the procedure. For example, while $7x-(3x-2)=38$ might be solved in four steps ($7x-(3x-2)=38 \rightarrow 7x-3x+2=38 \rightarrow 4x+2=38 \rightarrow 4x=36 \rightarrow x=9$), the solution expands to 20 steps as shown in Table 7.3. The large number of steps suggests that solving equations can involve a high number of operations that is routine if the learner has the requisite knowledge. In addition, use of a single mathematical operation per step allowed us to standardize the analysis.

When the process of solving $7x-(3x-2)=38$ is examined using O'Neil et al.'s cognitive readiness framework, the complexity of equation solving becomes clearer. The cognitive processes of adaptability and metacognition are particularly relevant.

Adaptability is defined as an effective change in response to an altered situation (O'Neil et al., this volume). A student competent in solving equations can respond effectively to different equation forms that involve multiple terms, grouping symbols, number types, operations, and properties, as initially presented to the learner as well as during the solution process whereby the form of the equation changes as the equation goes through successive transformations (or simplifications). This competency captures the fluency and automaticity identified as critical to the preparation for algebra by the National Mathematics Advisory Panel (NMAP, 2008). As students attain fluency with the symbols and operations required to transform equations, they may advance toward *adaptive expertise*—the possession of a deep understanding of the conceptual structure of the problem and understanding when and why particular procedures are appropriate or not. Students who have attained this stage are capable of solving equations across various surface forms (e.g., with the variable on both sides of the equal sign), are facile at iteratively simplifying complex expressions (e.g., nested quantities), and are able to recognize the optimal point during a solution path to eliminate terms and factors. Finally, successful students presumably engage in *metacognition*—composed of planning and self-monitoring—whereby the correctness of each solution step is monitored and the "chunk" size of a step is adjusted to reduce cognitive load and decrease errors.

7.3 Implications for an Assessment of Cognitive Readiness

An assessment of cognitive readiness for solving equations requires a way of measuring the process of solving equations to determine whether the student can respond appropriately to various forms of an equation that result from successive transformations. A key innovation we developed was to use the steps in the solution path of a given equation as a source to sample items from. A step in the solution path is treated as a test item for the participant to solve. By sampling a range of steps from a given equation, the item set inherently captures the *process* of solving that equation. The complexity of the items systematically decreases because solving an equation results in successive transformations of that equation into simpler equations.

Further, because all the steps flow from the same equation, the steps are internally coherent. Knowing the transitions from one step to the next—the transformation of each step into a simpler equation—is the key competency in solving equations. Solving an equation requires the learner to iteratively identify and execute the appropriate operation given evolving constraints of the equation. By testing students on each step, the assessment can identify where in the solution path students may be having difficulties.

The second innovation was inspired by Kalyuga and colleagues (Kalyuga, 2006; Kalyuga & Sweller, 2004). Their research has suggested that asking participants only for the next step (vs. a fully worked solution) is predictive of participants' performance on a fully worked solution (Kalyuga, 2006; Kalyuga & Sweller, 2004). We adopted this procedure as it is highly efficient. Thus, our research question was to what extent can steps in a solution (as in Table 7.3), when sampled as an assessment, capture the cognitive complexity of solving the equation they were derived from?

7.4 Method

7.4.1 Participants

Data for 42 participants were analyzed. The sample was from a larger study (Chung et al., 2007). Students were from an urban middle school in southern California. Students were tested at the end of the first semester. Participants were drawn from two sixth-grade algebra readiness (pre-algebra topics) classes and three eighth-grade algebra 1A classes (pre-algebra and algebra topics). There were 23 males and 17 females, and two participants did not report their sex. The students' ethnicity was diverse, including 19 % Latino, 33 % Asian or Pacific Islander, 24 % White, and 12 % African American, and 11 % unreported. About 80 % of students reported receiving A's or B's in math, and nearly all students agreed or strongly agreed that they were able to understand their teacher's explanations in math class, and nearly all students agreed or strongly agreed that they were able to read and understand most of the problems and explanations in their math textbook.

7.4.2 Measures

Pretest. A 27-item selected-response measure was used to measure students' knowledge and skills required to solve $7x-(3x-2)=38$, as described in Table 7.3. Detailed information on the measure is reported in Chung et al. (2007). Cronbach's α for this measure was 0.75.

Table 7.4 Next step items

Item	Equation	Step from Table 7.3
1	$7h-(3h-2)=38$	Given
2	$7x+(-1)(3x-2)=38$	2
3	$7a+(-3a)+2=38$	6
4	$2y+4=42$	11
5	$8x+6-6=24-6$	12
6	$9a+0=42-6$	13
7	$4x=38-2$	14
8	$5z=30$	15

Each item was based on a step in Table 7.3

Next step scale. An eight-item measure was developed to measure students' skills related to solving $7x-(3x-2)=38$. Each item was drawn from a step in the solution path shown in Table 7.3. For the next step items, participants were instructed to write down just their next step and not work through the full solution. The items were scored as correct or incorrect. The presentation order of the items was randomized on the form. The letters denoting variables were changed across the items and the values of terms and coefficients were changed so as to give the illusion of different equations. However, the structure of the equation was preserved. Cronbach's α for this scale was 0.69. Table 7.4 lists the items. Note that Table 7.4 also represents the form of the steps in a fully worked solution to $7x-(3x-2)=38$. In addition, substituting the appropriate value for each of the coefficients, terms, and variable labels in each item in Table 7.4 would yield the exact solution to $7x-(3x-2)=38$, which is the innovation of our general approach.

7.4.3 Task and Procedure

Participants were administered the measures as part of their normal math instruction. Participants completed the pretest, next step measure, and a background questionnaire. Participants were allowed the entire class period of 50 min to complete the tasks.

7.5 Results

Our research question focused on the extent to which steps from the solution path of an equation capture the cognitive complexity of solving the equation they were derived from. To address this question, our analysis examined (a) the extent to which an item's difficulty corresponded to its relative position in the solution path; and (b) the extent to which performance on the next step format predicted performance on a fully worked solution.

Table 7.5 Descriptive statistics and intercorrelations (Spearman) on background variables

Variable	*M*	SD	Min.	Max.	1	2
1. Next step assessment[a]	4.00	2.09	0	8	–	
2. Pretest[b]	20.60	3.46	12	27	0.44**	–
3. Self-reported grades in math[c]	1.59	0.84	1	4	–0.34*	–0.64***

$^{*}p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$

[a]Maximum score possible = 8

[b]Maximum score possible = 27

[c]1 = mostly A's, 2 = mostly B's, 3 = mostly C's, 4 = mostly D's, 5 = mostly F's

7.5.1 *Descriptive Statistics*

Means, standard deviations, and intercorrelations among knowledge of pre-algebra concepts, and self-reported math grades are shown in Table 7.5. Performance on the next step assessment, pretest, and self-reported grades were significantly correlated with each other. The mean performance on the next step assessment was 50 %, and 76 % for the pretest.

7.5.2 *To What Extent Does the Difficulty of an Item Correspond to Its Relative Position in the Solution Path?*

To answer this question, we examined the overall performance on each item relative to its position in the solution path to the equation $7x-(3x-2)=38$. Because solving an equation by definition results in a series of simpler and simpler equations, we expected an increase in performance across items that corresponded to simpler steps in the solution path.

The lowest item difficulty (or *p* value) was 0.17 for the most complex equation, and generally increased as the items became simpler. The highest *p* values were in the 0.7 range and were for the three simplest equations. Figure 7.1 shows the *p* values of each item. Higher item numbers indicate simpler equations. As expected, overall performance generally increased across simpler equations. In addition, the items appeared to cluster into three performance levels. Cluster 1 was the hardest and composed of the first two items in the solution path that were complex multistep equations where distribution was required to simplify the equations. Cluster 2 was composed of three items of multistep equations and did not require distribution. Cluster 3 was composed of items that were the last three steps in the solution path and were items that were generally single-step equations.

Because the items were dichotomous and dependent, a nonparametric procedure was performed to test whether there were differences in student performance on the eight items (Cochran's *Q* test, Pett, 1997). Cochran's *Q* test yielded a significant omnibus effect ($Q=67.34$, df $=1$, $p<0.001$), indicating a significant difference between at least two items. Follow-up multiple comparisons using Cochran's *Q* test

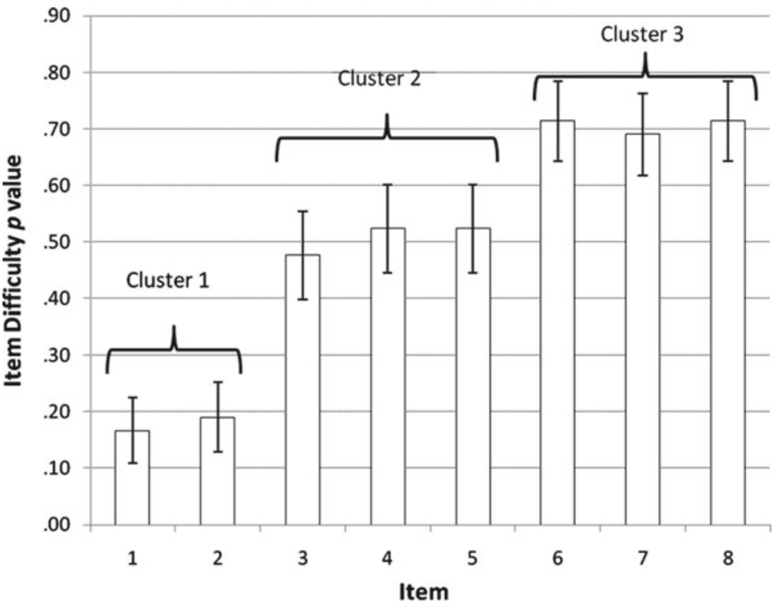


Fig. 7.1 *p* value of items corresponding to steps in the solution. Error bars indicate standard errors around the mean

which were conducted between items that corresponded to adjacent steps in the solution (e.g., 1 vs. 2; 2 vs. 3) showed significant differences between items 2 and 3 ($Q = 10.29$, df = 1, $n = 42$, $p = 0.001$) and items 5 and 6 ($Q = 4.57$, df = 1, $n = 42$, $p = 0.03$). No other differences were detected, suggesting that the difficulty of the items within each item cluster was similar.

An inspection of the items suggested that items within each cluster reflected similar cognitive demands. Cognitive demands refer to the type of cognitive processing required of the student to be successful on an assessment task (Baker, 1997). In this case, the major next step for items 1 and 2 is distribution over subtraction. For items 3–5, the major next step is subtracting terms from both sides of the equation (addition property of equality), and for items 6–8 the major next step is dividing both sides of the equation (multiplication property of equality).

Items within each cluster were summed to form a scale as student performance on the items within a scale was similar and the math operations were mainly simplification. Interestingly, Cluster 1 (distribution) was related significantly to Cluster 2 (addition property of equality), $r_s(41) = 0.39$, $p = 0.01$, but not to Cluster 3 (multiplication property of equality), while Cluster 2 was related significantly to Cluster 3, $r_s(41) = 0.32$, $p = 0.04$. These relationships are consistent with how the cognitive demands unfold during solving the equation. The major next step after Cluster 1 is Cluster 2, and the major next step from Cluster 2 is Cluster 3. In terms of cognitive readiness, the discontinuities after Cluster 1 suggest student difficulties with distribution over subtraction. The discontinuity after Cluster 2 suggests

Table 7.6 Comparison between the next step only and fully worked solution items ($N=42$)

	Next step response		Fully worked		
Item pair	Equation	Item difficulty[a]	Equation	Item difficulty[a]	*p* value for test of item differences[b]
1	$5z=30$	0.71	$6m=42$	0.98	0.001
2	$2y+4=42$	0.52	$3g+1=13$	0.79	0.003
3	$7h-(3h-2)=38$	0.17	$7h-(3h-2)=38$	0.14	0.66

[a]Proportion of students who answered correctly
[b]Cochran's test

some students were having difficulty with the transformation step—reasoning about when and how to "eliminate" terms and coefficients in an expression to simplify the equation.

7.5.3 To What Extent Does Asking for Just the "Next Step" Predict Performance on a Fully Worked Solution?

To answer this question, we examined the relation between participants' performance on three next step items and their performance on three similar items that required participants to solve the items completely. Our analysis approach first examined format differences between the next step and fully worked items. Because we were examining whether the next step format could be used as a substitute for the typical approach of requiring students to solve each equation completely, we examined whether there existed format differences and whether the next step items predicted performance on the fully worked items. Both of these properties would be required if the next step item format were to be used to measure students' skill at solving equations.

Table 7.6 shows the items and item statistics for the next step and fully worked item formats. The next step format appeared to underestimate students' skill at actually solving the equation. Cochran's test of whether there were differences in item types yielded a significant difference for item pair 1 ($Q=11.0$, $\mathrm{df}=1$, $p=0.001$) and item pair 2 ($Q=9.0$, $\mathrm{df}=1$, $p=0.003$). There was no difference between the formats for item pair 3. These results point to a potential format effect. The task to write down just the next step and not provide the full solution may have been unusual. Another difference was that the values of the coefficients were different across formats, which may have contributed to differences in difficulty.

The next analysis examined how well performance on the next step item predicted performance on the fully worked item. These analyses could not be performed for item pair 1 because all participants answered the item correctly on the fully worked format. For item pair 2, the prediction of performance on the fully worked problem from the next step performance was marginally significant (Somer's $d=0.21$, $p=0.09$), and significant for item pair 3 (Somer's $d=0.51$, $p=0.03$). Somer's d is a measure of concordance between two ordinal variables and ranges in value from −1 to +1. Values close to −1 or +1 indicate a strong negative or positive relationship, respectively, and values close to 0 indicate little or no relationship between the variables (IBM SPSS, 2009).

We then summed items within each format to form scales and examined the correlation between the two scales. The purpose for forming a scale was to create a more general measure that spanned the full range of solving equation steps. The next step scale correlated significantly with the fully worked scale, $r_s(40)=0.39$, $p=0.012$, although the magnitude of the correlation was lower than that reported in other research (rs between 0.7 and 0.9, Kalyuga, 2006; Kalyuga & Sweller, 2004). The difference in magnitude may be due to the small number of items in each scale. These results, while based on a small sample, suggest that the next step item format can be predictive of whether participants will be successful at solving an equation.

7.6 Discussion

In this study we tested a novel assessment technique to measure the cognitive readiness for solving equations. Our technique, inspired from Kalyuga and colleagues (Kalyuga, 2006; Kalyuga & Sweller, 2004), sampled steps from the solution path of an equation and used those steps as assessment items. We also examined whether simply asking participants to specify their "next step" captured the complexity of solving the entire equation.

Our first finding was that items drawn from steps from a solution path yielded item difficulties consistent with the step's relative position in the solution path. Items drawn from the beginning of the solution path were more difficult than items near the end of the solution path. However, item difficulties were similar among items that differed only in the arithmetic complexity (e.g., simple subtraction or division). The major discontinuities in performance occurred in the steps that required operations related to distribution and equality (subtraction and division). This result is consistent with the finding that many students have neither the skills nor precise understanding of the body of basic mathematical knowledge to successfully transform equations (e.g., Demby, 1997; Herscovics & Linchevski, 1994; Kieran, 2007; MacGregor & Stacey, 1997; Pierce & Stacey, 2007).

Our second finding was that the next step item format was apparently more difficult than like items that required participants to work out the full solution, which was similar to a finding by Kalyuga (2006). Kalyuga found that students performed lower on the next step items than on items requiring fully worked solutions. However, Kalyuga used word problems as the task and overall performance was low for both formats. Kalyuga speculated that students' problem solving skills were impoverished which led to a lower success rate on the next step items because providing an accurate next step required an existing schema of the general approach, compared to solving the problem where the solution could be discovered through a variety of approaches. These results, however, were inconsistent with earlier work using a similar approach. In Study 1 that used an equation solving task, Kalyuga and

Sweller (2004) found that students solved a higher percentage of the next step items (72 %) than the fully worked items (58 %). While the general trend of format differences appears to exist among all the studies, which format is more difficult is unclear as all studies scored student responses differently. We think the most likely explanation for the differences observed in the current study is that the next step format is too novel, as asking students to write only their first step is atypical. However, despite these format differences, performance on next step items predicted performance on items requiring a fully worked solution in the current study as well as Kalyuga (2006) and Kalyuga and Sweller (2004).

Given the amount of knowledge and skills required of solving equations, an important assessment question is what should be measured and how should it be measured? The use of assessments as diagnostic tools is not new and has seen numerous forms, many of which are clinical and intensive in nature (Black & Wiliam, 2009; Heritage, Kim, Vendlinski, & Herman, 2009; Sadler, 1989; Shepard, 2001; Wiliam & Thompson, 2007).

The objective of an assessment is to understand what the student is doing and to elicit why he or she is doing it. The cognitive demands of solving equations include knowing which operation to apply (e.g., distribution over subtraction) and, just as importantly, knowing when to apply those operations (e.g., equality property of addition). In the context of cognitive readiness, the assessment presented in this chapter was designed to measure both the prerequisites for solving equations and the skill itself. In addition, the technique may provide a feasible way to capture the *process* of solving equations. Our results show that the prerequisites for solving equations can be measured feasibly. Our general approach of sampling from the solution path yields performance differences that suggest chokepoints that in the context of solving equations map to adaptive expertise and adaptability. That is, successful problem solvers appear to know the conditions under which to apply particular operations to solve the equation (e.g., using the additive identity to isolate terms; using the multiplicative identity to isolate variables).

Finally, the practical use for algebra instruction is straightforward: because the item set is sampled from the derivation of an equation (e.g., $7x-(3x-2)=38$, as shown in Table 7.3), the performance dropoff can be used to pinpoint where in the solution path students have difficulty. Further, the set of steps in Table 7.3 tap all the properties of algebra and most of the operations. The item clusters shown in Fig. 7.1 suggest that, in our sample, single-step problems requiring division of the coefficient to isolate the unknown were relatively easy compared to multistep equations that required the use of the additive inverse. The very low performance on the first item cluster suggests that distribution over subtraction is posing a substantial barrier for students. The instructional implications of Fig. 7.1 are clear—students in our sample need support on (a) the use of the additive inverse to isolate the term with the unknown and (b) distribution over subtraction. These implications would not be as straightforward if the items were developed by other means. Our technique of sampling items from the set of steps in a solution path appears to be a promising approach, combining rapid testing time, breadth of coverage, and diagnostic potential.

Acknowledgments The work reported herein was supported under the National Research and Development Centers Program, PR/Award Number R305C080015, as administered by the Institute of Education Sciences, U.S. Department of Education, and partially supported by a grant from the Office of Naval Research, Award Number N000140810126. The findings and opinions expressed in this chapter do not necessarily reflect the positions or policies of the U.S. Department of Education or the Office of Naval Research. We would also like to thank Joanne Michiuye of UCLA/CRESST for review and editorial help with this manuscript.

References

Baker, E. L. (1997). Model-based performance assessment. *Theory Into Practice, 36*, 247–254.

Black, P., & Wiliam, D. (2009). Developing the theory of formative assessment. *Educational Assessment, Evaluation and Accountability, 21*, 5–31.

California Department of Education (CDE). (2011). *California High School Exit Exam (CAHSEE) result for Los Angeles Unified School District, 2010–2011 school year*. Sacramento, CA: Author.

California State University (CSU). (2011). *Fall 2010 final regularly admitted first-time freshman remediation systemwide*. Long Beach, CA: Author.

Carraher, D. W., & Schliemann, A. D. (2007). Early algebra and algebraic reasoning. In F. K. Lester Jr. (Ed.), *Second handbook of research on mathematics teaching and learning* (2nd ed., pp. 669–704). Charlotte, NC: Information Age Publishing.

Chazan, D., & Yerushalmy, M. (2003). On appreciating the cognitive complexity of school algebra: Research on algebra learning and directions of curricular change. In J. Kilpatrick, W. G. Martin, & D. Schifter (Eds.), *A research companion to principles and standards for school mathematics* (pp. 123–135). Reston, VA: National Council of Teachers of Mathematics.

Chung, G. K. W. K., Delacruz, G. C., Dionne, G. B., Baker, E. L., Lee, J. J., & Osmundson, E. (2007). Towards individualized instruction with technology-enabled tools and methods. In R. Perez (Chair), *Rebooting the past: Leveraging advances in assessment, instruction, and technology to individualize instruction and learning.* Symposium presented at the annual meeting of the American Educational Research Association, Chicago, IL.

Demby, A. (1997). Algebraic procedures used by 13–15 year-olds. *Educational Studies in Mathematics, 33*, 45–70.

Heritage, M., Kim, J., Vendlinski, T., & Herman, J. (2009). From evidence to action: A seamless process in formative assessment? *Educational Measurement: Issues and Practice, 28*, 24–31.

Herscovics, N., & Linchevski, L. (1994). A cognitive gap between arithmetic and algebra. *Educational Studies in Mathematics, 27*, 59–78.

IBM SPSS. (2009). *PASW Statistics (version 18) [Computer software]*. Armonk, NY: IBM.

Kalyuga, S. (2006). Rapid cognitive assessment of learners' knowledge structures. *Learning and Instruction, 16*, 1–16.

Kalyuga, S., & Sweller, J. (2004). Measuring knowledge to optimize cognitive load factors during instruction. *Journal of Educational Psychology, 96*, 558–568.

Kieran, C. (2007). Learning and teaching algebra at the middle school through college levels. In F. K. Lester Jr. (Ed.), *Second handbook of research on mathematics teaching and learning* (2nd ed., pp. 707–762). Charlotte, NC: Information Age Publishing.

MacGregor, M., & Stacey, K. (1997). Students' understanding of algebraic notation: 11–15. *Educational Studies in Mathematics, 33*, 1–19.

National Academy of Engineering (NAE), & Institute of Medicine (IOM), Committee on Science, Engineering, and Public Policy. (2006). *Rising above the gathering storm: Energizing and employing America for a brighter economic future*. Washington, DC: National Academies Press.

National Academy of Sciences (NAS). (2005). *National science and technology strategies in a global context: Report of an international symposium*. Washington, DC: National Academies Press.

National Center for Education Statistics (NCES). (2010). *The nation's report card: Grade 12 reading and mathematics 2009 national and pilot state results*. Washington, DC: U.S. Department of Education.

National Mathematics Advisory Panel (NMAP). (2008). *Foundations for success: The final report of the National Mathematics Advisory Panel*. Washington, DC: U.S. Department of Education.

National Science Board. (2010). *Science and engineering indicators, Volume 1* (Rep. No. NSB 10-02). Arlington, VA: Author.

Pett, M. A. (1997). *Nonparametric statistics for health care research: Statistics for small samples and unusual distributions*. Thousand Oaks, CA: Sage.

Pierce, R., & Stacey, K. (2007). Developing algebraic insight. *Mathematics Teaching, 203*, 12–16.

Sadler, D. R. (1989). Formative assessment and the design of instructional systems. *Instructional Science, 18*, 119–144.

Shepard, L. A. (2001). The role of classroom assessment in teaching and learning. In V. Richardson (Ed.), *The handbook of research on teaching* (4th ed.). Washington, DC: American Educational Research Association.

Vlassis, J. (2002). Making sense of the minus sign or becoming flexible with "negativity". *Learning and Instruction, 14*, 469–484.

Wiliam, D., & Thompson, M. (2007). Integrating assessment with instruction: What will it take to make it work? In C. A. Dwyer (Ed.), *The future of assessment: Shaping teaching and learning* (pp. 53–82). Mahwah, NJ: Erlbaum.

Chapter 8
What Problem Solvers Know: Cognitive Readiness for Adaptive Problem Solving

Richard E. Mayer

8.1 Introduction

Adaptive problem solving involves the ability to invent solutions to problems that the problem solver has not encountered before. In adaptive problem solving, problem solvers must adapt their existing knowledge to fit the requirements of a novel problem, including a novel problem situation that may change over time. Mayer and Wittrock (2006) refer to this situation as one that requires *problem-solving transfer*.

Consider the problem shown in Fig. 8.1. You are shown a diagram depicting a car's braking system and asked to troubleshoot what could cause it to fail. For people who are not expert car mechanics, answering this question requires adaptive problem solving. For example, some acceptable answers are that there may be a leak in the brake tube or there is too much space between the piston and the cylinder. Even after receiving a multimedia presentation on how a car's braking system works, most novices have difficulty in solving the brake problem (Mayer, 2009).

Next, consider the problem shown in Fig. 8.2. This problem comes in level 10 (i.e., an embedded transfer test) of an educational game intended to teach how electrical circuits work (Johnson & Mayer, 2010; Mayer & Johnson, 2010). Again, for people who are not experts in electronics, this problem requires adaptive problem solving. The correct answer is "same." Even after playing through nine levels of the game intended to teach basic principles of electric flow in circuits, a substantial proportion of players give the wrong answer.

Finally, consider the problem shown in Fig. 8.3. You are shown a diagram depicting part of the process of a solar cell and asked to troubleshoot what could go wrong. If you are not an expert in this field, this is an adaptive problem-solving situation for you. For example, some acceptable answers are that there are no free

R.E. Mayer (✉)
Department of Psychology, University of California, Santa Barbara,
Santa Barbara, CA 93106, USA
e-mail: mayer@psych.ucsb.edu

H.F. O'Neil et al. (eds.), *Teaching and Measuring Cognitive Readiness*,
DOI 10.1007/978-1-4614-7579-8_8,

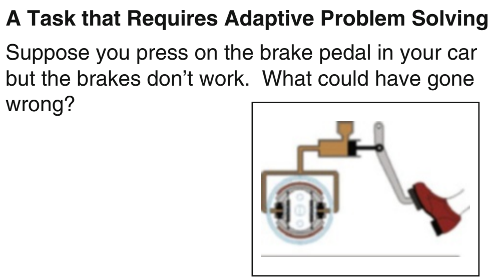


Fig. 8.1 The brake problem

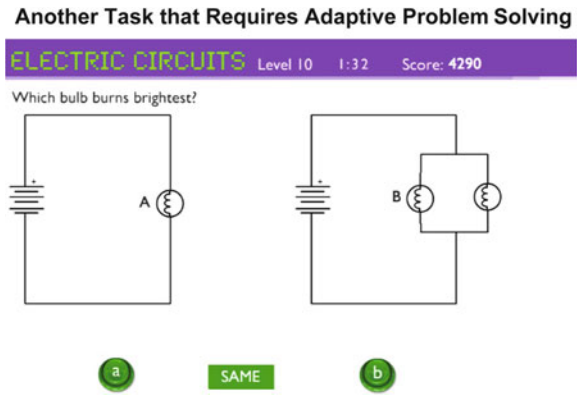


Fig. 8.2 The bulb problem

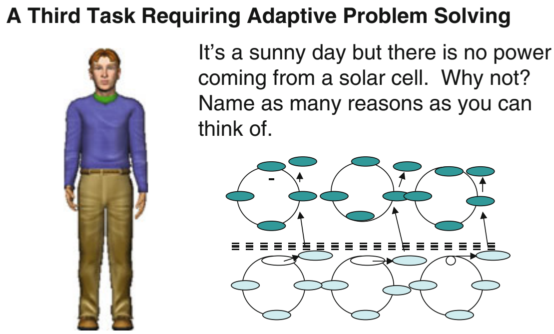


Fig. 8.3 The solar cell problem

electrons on the top layer or no free bonding sites on the bottom layer. After viewing a PowerPoint presentation delivered by an onscreen agent, most learners in our studies still have much difficulty in solving this problem (Mayer & DaPra, 2012).

In short, Figs. 8.1–8.3 represent three examples of adaptive problem solving based on research studies in our lab at the University of California, Santa Barbara. In each example of adaptive problem solving, the problem solver must adapt to a new situation that he or she has not encountered before. Preparing people to be ready to engage in adaptive problem solving constitutes a major challenge for

Table 8.1 Three components in cognitive readiness for adaptive problem solving

Component	Issue
Science of learning	What do you need to know to be ready for adaptive problem solving?
Science of assessment	How can we know what you know?
Science of instruction	How can we help you learn what you need to know to be ready for adaptive problem solving?

trainers and trainees. In short, cognitive readiness for adaptive problem solving refers to being ready to solve novel problems.

In this chapter, I examine three issues concerning cognitive readiness for adaptive problem solving as summarized in Table 8.1. First, based on the science of learning, an important issue concerns specifying what someone needs to know in order to be ready for adaptive problem solving. Second, based on the science of assessment, an important issue concerns how to determine what someone knows. Third, based on the science of instruction, an important issue concerns discovering effective instructional methods that help people learn what they need to know to be ready for adaptive problem solving. As you can see, the thesis of this chapter is that knowledge is at the heart of cognitive readiness—including how to specify the needed knowledge, how to assess whether learners have it, and how to help learners create it. I examine these three issues in the next three sections of the chapter, respectively.

8.2 Science of Learning: What Is Cognitive Readiness for Adaptive Problem Solving?

8.2.1 Role of Knowledge in Adaptive Problem Solving

The science of learning is the scientific study of how people learn. Learning is a change in the learner's knowledge. An important contribution of the science of learning involves techniques for specifying the desired knowledge change—including the knowledge needed to support adaptive problem solving.

Cognitive readiness for adaptive problem solving depends on the learner's knowledge. Table 8.2 lists five kinds of knowledge that support cognitive readiness for adaptive problem solving—facts, concepts, procedures, strategies, and beliefs (Anderson et al., 2001; Mayer, 2008). People may need to integrate all five kinds of knowledge in order to perform well on tasks requiring adaptive problem solving. Strategies include metacognitive strategies for monitoring and controlling one's cognitive processing on a problem-solving task. Beliefs can affect the learner's motivation to initiate and maintain efforts to solve a problem.

As you see, the learner's knowledge plays an essential role in adaptive problem solving. I use knowledge in the broad sense to include all five kinds of knowledge listed in Table 8.2. This analysis is equivalent to the classic distinction among knowledge (which corresponds to facts and concepts), skills (which corresponds to procedures and strategies), and attitudes (which corresponds to beliefs).

Table 8.2 Five kinds of knowledge that support cognitive readiness for adaptive problem solving

Kind	Description	Example
Facts	Factual knowledge about a domain	Brake fluid contains ethylene glycol
Concepts	Categories, schemas, models, principles	A piston in a cylinder is like a syringe
Procedures	Step-by-step processes	Conversion of 1/2 in. to millimeters
Strategies	General methods	Break a problem into parts
Beliefs	Thoughts about one's cognition	"I am good at working with mechanical devices"

8.2.2 A Closer Look at Adaptive Problem Solving

In order to specify the knowledge required for readiness for adaptive problem solving, it is useful to clarify the nature of problems and problem solving.

What is a problem? A problem exists when (a) a situation is a given state, (b) the problem solver's goal is to have the situation in a desired state, and (c) there is no obvious way to move from the given state to the desired state. The Gestalt psychologist Karl Duncker (1945, p. 1) eloquently summarized this classic definition as, "A problem exists when a living creature has a goal but does not know how the goal is to be reached."

It is customary to distinguish between routine and nonroutine problems. With *routine problems*, the problem solver knows a solution method. For example, for most adults, "What is 22 × 155?" is a routine problem because they know the procedure for long multiplication. With *nonroutine problems*, the problem solver does not immediately know a solution method and therefore must invent one. For example, for nonexperts, "How can we make a solar cell more effective?" is a nonroutine problem. As you can see, technically, routine problems do not meet the definition of a problem because "there is an obvious way to move from the given state to the goal state" (or to put it another way, the problem solver "knows how the goal is to be reached"). Thus, adaptive problem solving refers to solving nonroutine problems rather than routine problems.

It is also customary to distinguish between well-defined and ill-defined problems. A *well-defined problem* has a clear statement of the givens, goals, and allowable moves. Examples include solving arithmetic problems or playing a game of chess. An *ill-defined problem* does not have a clear statement of the givens, goals, and/or allowable operators. For example, "Solve the energy crisis" is an ill-defined problem because what you are allowed to do (i.e., the allowable moves) are not clearly specified, nor is the desired goal state. Adaptive problem solving generally refers to solving ill-defined rather than well-defined problems.

What is problem solving? Problem solving is directed cognitive processing. Problem is directed because it seeks to achieve a goal. Problem solving is cognitive because it occurs in the problem solver's mind and must be inferred indirectly through the problem solver's behavior. Problem solving is processing because it involves the mental manipulation of knowledge representations in the problem solver's cognitive system. The Gestalt psychologist Karl Duncker (1945, p. 1) eloquently captured this definition in his classic description, "Whenever one cannot go from the given state to the desired state simply by action, then there is recourse to thinking. Such thinking has the task of devising some action which may mediate between the existing and desired situations." As you can see, figuring out how to accomplish a goal is the essence of problem solving.

How does problem solving work? Problem solving involves two major phases—*problem representation*, in which the problem solver constructs a mental representation of the problem—and *problem solution*, in which the problem solver devises and carries out a solution plan. Although there are several conceptions of how problem solving works (Mayer, 1995), the most prominent explanation is that problem solving occurs when the problem solver reformulates the problem—that is, conceives of the givens or the goal in a new way. The Gestalt psychologist Karl Duncker (1945, p. 1) eloquently summarized this idea as, "What is really done in any solution of problems consists in formulating the problem more productively." Once the problem solver can represent the problem more productively, it is easier to arrive at a solution plan.

What is adaptive problem solving? In light of this analysis, I can expand on the definition of adaptive problem solving. Adaptive problem solving is a form of problem solving that requires a series of problem reformulations or continual reevaluation of problem formulations in light of changing conditions. In short, adaptive problem solving occurs when a problem solver continually revises how he or she represents the problem (and its solution plan) in light of the changes in the problem situation.

8.3 Science of Assessment: How Can We Measure Cognitive Readiness for Adaptive Problem Solving?

The science of assessment is the scientific study of how to determine what people know. For example, Table 8.3 presents example assessment items for each of the five kinds of knowledge required for cognitive readiness—facts, concepts, procedures, strategies, and beliefs.

Although these items may tap individual aspects of the required knowledge, it is also important to assess ways of using this knowledge that go beyond simply remembering. Table 8.4 summarizes six kinds of cognitive processes that may be tapped by assessment items—remember, understand, apply, analyze, evaluate, and

Table 8.3 Example assessment items for five kinds of knowledge required for cognitive readiness

Kind	Example
Facts	In Ohm's law, R stands for ___________
Concepts	What would happen to the rate of current if we added another resistor in parallel?
Procedures	If $V=10$ and $R=2$, compute the value of I
Strategies	In a circuit with a current of 5 A, judge which of several different methods accomplishes the goal of doubling the current
Beliefs	Rate from 1 (strongly disagree) to 5 (strongly agree): I like playing the Ohm's Law Game

Table 8.4 Six kinds of cognitive processes in using knowledge

Kind	Description	Example
Remember	Retrieve relevant knowledge from long-term memory	State the formula for Ohm's law
Understand	Construct meaning from instructional messages	Restate Ohm's law in your own words
Apply	Carry out a procedure in a given situation	Compute the value of *I* given *V* and *R*
Analyze	Break materials into parts; determine how parts relate	Distinguish relevant numbers in a word problem
Evaluate	Make judgments based on criteria	Determine the best way to solve a word problem
Create	Put elements together to form a coherent whole	Plan an essay on the history of Ohm's law

Table 8.5 Four types of transfer questions

Type	Example
Troubleshooting	Suppose you pull up and push down on the handle of a bicycle tire pump but no air comes out. What could have gone wrong?
Redesign	How could you make brakes more effective—that is, how could you reduce the distance needed to stop?
Prediction	What would happen if you reversed the positive and negative wires on an electric motor?
Explanation	What does temperature have to do with lightning formation?

create. This taxonomy is based on a revision of the Bloom's taxonomy of educational objectives (Anderson et al., 2001). As you can see, the latter five processes involve using the learned material in new ways rather than simply remembering it—and thus offer possible techniques for assessing cognitive readiness.

Two ways to measure learning outcomes are retention tests and transfer tests. *Retention tests* focus on remembering by asking the learner to recall or recognize the presented material. For example, after reading a lesson on how a car's braking system works, a learner may be asked, "Write down all you can remember about how a car's braking system works as described in the lesson." *Transfer tests* focus on understanding by asking the learner to evaluate or use the material in a new situation. For example, after a lesson on how a car's braking system works, a learner may be asked, "How would you improve the effectiveness of the braking system described in the lesson?"

The pattern of performance on retention and transfer tests indicates three kinds of learning outcomes (Mayer, 2008, 2009). No learning is indicated by poor performance on retention and transfer tests. Rote learning is indicated by good performance on retention and poor performance on transfer. Meaningful learning is indicated by good performance on retention and transfer tests. Thus, cognitive readiness for adaptive problem solving is evidenced by problem-solving transfer in addition to simply focusing on retention.

As you can see, transfer tests are essential in assessment for cognitive readiness for adaptive problem solving. Table 8.5 provides examples of transfer items that require using the information from a lesson in new situations. In particular, the transfer items involve troubleshooting, redesign, prediction, and explanation of a cause-and-effect system (such as how a pump works, how an electric motor works, how an electric circuit works, or how a solar cell works). Transfer items are intended to require adaptive problem solving.

8.4 Science of Instruction: How Can We Teach Cognitive Readiness for Adaptive Problem Solving?

8.4.1 Which Cognitive Processes During Learning Promote Cognitive Readiness?

The science of instruction is the scientific study of how to help people learn. In particular, effective instruction works because it encourages the learner to engage in appropriate cognitive processing during learning. To understand how to accomplish this goal, let's begin with the cognitive model of multimedia learning as presented in Fig. 8.4. This model reflects three principles of how the human cognitive system works based on research in cognitive science:

- *Dual channels principle*—People have separate channels for processing visual/pictorial and auditory/verbal information.
- *Limited capacity principle*—People can engage in only a small amount of cognitive processing in each channel in working memory at any one time.
- *Active processing principle*—Meaningful learning occurs when people engage in appropriate cognitive processing during learning.

The boxes in Fig. 8.4 represent three memory stores:

- *Sensory memory*—which holds visual and auditory images for a brief time (e.g., less than a quarter of a second) with unlimited capacity
- *Working memory*—which can store and manipulate only a few visual items and a few verbal items at any one time (so they are lost within a half minute if they are not processed)
- *Long-term memory*—which permanently holds a storehouse of organized knowledge with unlimited capacity

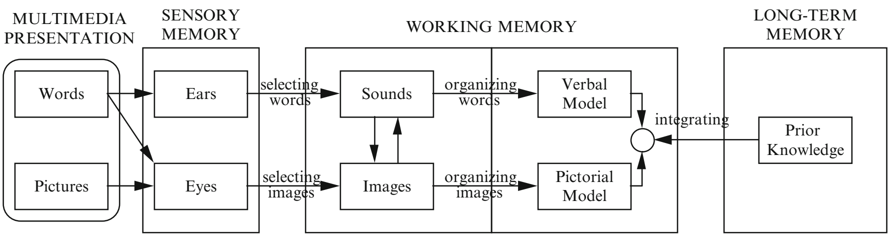


Fig. 8.4 Cognitive theory of multimedia learning

Table 8.6 Three demands on cognitive capacity

Demand	Description	Caused by
Extraneous processing	Cognitive processing that does not support the objective of the lesson	Poor instructional design
Essential processing	Cognitive processing required to mentally represent the presented material	Complexity of the content
Generative processing	Deep cognitive processing aimed at making sense of the presented material	Learner's motivation to exert effort

The arrows in Fig. 8.4 highlight three kinds of cognitive processing during learning:

- *Selecting*—in which the learner attends to relevant aspects of the incoming visual and auditory information (as indicated by the arrows from sensory memory to working memory)
- *Organizing*—in which the learner mentally arranges the pieces of information in working memory into coherent verbal and pictorial representations (as indicated by the arrows within working memory)
- *Integrating*—in which the learner connects the verbal and pictorial representations with each other and with relevant prior knowledge activated from long-term memory (as indicated by the arrow from long-term memory to working memory)

Cognitive readiness for adaptive problem solving is promoted by instructional methods that guide the learner's appropriate cognitive processes during learning—including selecting relevant words and pictures from the presented material, organizing them into coherent verbal and pictorial representations, and integrating the representations with each other and with knowledge activated from long-term memory. An important constraint is that working memory is limited in capacity so only a limited amount of cognitive processing can take place at any one time.

Table 8.6 describes three demands on the learner's cognitive capacity during learning—extraneous processing (which is wasted processing that does not address the instructional objective), essential processing (which consists of selecting and initial organizing), and generative processing (which consists of extensive organizing and integrating). Instruction for cognitive readiness for adaptive problem solving has the challenge of fostering generative processing while at the same time managing essential processing and minimizing extraneous processing.

8.4.2 What Works in Improving Problem-Solving Transfer?

An important accomplishment of the science of instruction for cognitive readiness is a set of evidence-based principles for how to design effective instruction—that is instruction that enables problem-solving transfer. Evidence-based practice refers to the idea that instructional methods should be based on empirical research evidence,

Table 8.7 Five evidence-based principles for reducing extraneous processing

Principle	Description	Example
Coherence [1,2,4]	People learn better when extraneous material is excluded rather than included	Cut out interesting but irrelevant text and graphics
Signaling [1,4]	People learn better when the organization of a lesson is highlighted	Use outlines and section headings for a text lesson
Spatial contiguity [1,2,3,4]	People learn better when corresponding printed words and pictures are near rather than far from each other on the screen or page	Embed relevant words within a graphic rather than as a caption
Temporal contiguity [1,2,4]	People learn better when corresponding spoken words and pictures are presented simultaneously rather than successively	Present narration at the same time as animation rather than before or after
Expectation [2]	People learn better when they are shown the type of test items they will receive following learning	Tell people that after reading this chapter, they will be asked to tell how to define adaptive problem solving

including experiments that compare learning outcomes of students who learn with vs. without a particular instructional feature. Summaries of evidence-based principles include:

1. Mayer's (2005) *Cambridge handbook of multimedia learning*
2. Halpern, Graesser, and Hakel's (2007) *25 learning principles to guide pedagogy and the design of learning environments*
3. Pashler et al.'s (2007) *Organizing instruction and study to improve student learning*
4. O'Neil's (2005) *What works in distance learning: guidelines*

An extraneous overload situation occurs when poor instructional design creates the need for so much extraneous processing that the learner has inadequate remaining cognitive capacity to engage in needed essential and generative processing. Table 8.7 describes five evidence-based principles for reducing extraneous processing—coherence, signaling, spatial contiguity, temporal contiguity, and expectation. The numbers in brackets correspond to the four sources of evidence-based principles listed in the preceding paragraph. As you can see in Table 8.7, learners are less likely to engage in extraneous processing when extraneous words and pictures are eliminated from a lesson (coherence principle), when essential words and pictures are highlighted in a lesson (signaling principle), when corresponding printed words and graphics are near each other on the page or screen (spatial contiguity principle), when corresponding narration and graphics are presented at the same time (temporal contiguity principle), and when learners are told in advance about what to expect for the test (expectation principle).

An essential overload situation occurs when the content of lesson is so complex that the learner lacks enough cognitive capacity to engage in the required essential

Table 8.8 Three evidence-based principles for managing essential processing

Principle	Description	Example
Segmenting [1,2,4]	People learn better when a complex lesson is presented in manageable parts	Break a narrated animation into small segments each with a CONTINUE button
Pre-training [1,4]	People learn better from a complex lesson when they receive pre-training in the names and characteristics of the key concepts	Tell people about the names, locations, and characteristics of the parts before showing them a narrated animation
Modality [1,3,4]	People learn better from a multimedia presentation when words are spoken rather than printed	Accompany an animation with a spoken description rather than onscreen captions

Table 8.9 Four evidence-based principles for fostering generative processing

Principle	Description	Example
Multimedia [1,2,4]	People learn better from words and pictures rather than from words alone	Add relevant graphics to text
Personalization [1,4]	People learn better when words are in conversational style rather than formal style	Use "I" and "you"
Concretizing [2,3]	People learn better when unfamiliar material is related to familiar knowledge	Provide concrete examples or analogies
Anchoring [2,3]	People learn better when material is presented within the context of a familiar situation	Learn about functions within the context of a business financial plan

processing. Table 8.8 describes three evidence-based principles for managing essential processing, thereby freeing capacity for needed essential and generative processing—segmenting, pre-training, and modality. As you can see in Table 8.8, learners are better able to manage essential processing when complex material is broken into bite-sized segments and presented with a continue button (segmenting principle), when learners receive pre-training in the names, locations, and characteristics of key elements (pre-training principle), and when words in a multimedia presentation are presented in spoken form thereby offloading some processing from the visual channel to the verbal channel (modality principle).

An underutilization situation occurs when a learner has the cognitive capacity available for the required essential and generative processing, but the learner is not motivated to exert the effort to engage in this processing. Table 8.9 describes four evidence-based principles for fostering generative processing—multimedia, personalization, concretizing, and anchoring (Mayer, 2009, 2011). As you can see, learners work harder to make sense of the presented material when the lesson contains words and pictures rather than words alone (multimedia principle), when words are presented in conversation style such as using first and second person constructions (personalization principle), when abstract material is explicitly linked to concrete examples or analogies (concretizing principle), and when unfamiliar material is presented within a familiar context or scenario (anchoring principle).

Other evidence-based techniques that encourage generative processing include the testing principle (e.g., having the learner answer practice test items after studying a lesson), the self-explanation principle (e.g., asking the learner to point out anything that does not make sense while reading a lesson), the questioning principle (e.g., asking the learner to produce possible test questions after viewing a presentation), and the elaboration principle (e.g., asking the learner to take notes while listening to a lecture) (Halpern et al., 2007; Mayer, 2011; Pashler et al., 2007). Evidence-based techniques geared for practicing include the feedback principle (e.g., providing a step-by-step explanation for the correct answer after someone attempts to solve a problem), the worked examples principle (e.g., providing a step-by-step explanation of how to solve a problem before asking someone to solve a similar problem), and the guided discovery principle (e.g., asking someone to solve a problem while the instructor provides hints, models how to solve parts of the problem, and offers explanative feedback) (Halpern et al., 2007; Mayer, 2005, 2011; O'Neil, 2005; Pashler et al., 2007). As you can see, the science of instruction is enjoying some success in deriving principles for how to help people learn in ways that prepare them for adaptive problem solving as can be found in Clark and Mayer (2011).

8.5 Conclusion

In conclusion, based on the science of learning, we began with the idea that cognitive readiness for adaptive problem solving depends on the learner's knowledge (i.e., facts, concepts, procedures, strategies, and beliefs). Based on the science of assessment, we seek to assess adaptive problem solving by examining problem-solving transfer tests. Based on the science of instruction, we consider instructional methods that promote cognitive readiness for adaptive problem solving through priming appropriate cognitive processing during learning (e.g., multimedia, personalization, generating, anchoring, testing, self-explanation, and elaboration). In short, understanding cognitive readiness for adaptive problem solving involves specifying the knowledge that enables adaptive problem solving, determining how to assess whether a learner possesses that knowledge, and discovering how to help people acquire it. This line of research is an example of what it means to apply the science of learning to education (Mayer, 2011).

Acknowledgements Writing of this chapter was supported by grants N00014080018 and N000140810126 from the Office of Naval Research. The findings and opinions expressed here do not necessarily reflect the positions or policies of the Office of Naval Research.

References

Anderson, L. W., et al. (2001). *A taxonomy for learning, teaching, and assessing*. New York: Longman.

Clark, R. C., & Mayer, R. E. (2011). *e-Learning and the science of instruction* (3rd ed.). San Francisco: Pfeiffer.

Duncker, K. (1945). On problem solving. *Psychological monographs, 58*(5), Whole No. 270.
Halpern, D. F., Graesser, A., & Hakel, M. (2007). *25 learning principles to guide pedagogy and the design of learning environments.* Washington, DC: Association for Psychological Science Taskforce on Lifelong Learning at Work and Home. Retrieved from http://www.psyc.memphis.edu/learning/whatweknow/index.shtml
Johnson, C. I., & Mayer, R. E. (2010). Adding the self-explanation principle to multimedia learning in a computer-based game-like environment. *Computers in Human Behavior, 26*, 1246–1252.
Mayer, R. E. (1995). The search for insight: Grappling with Gestalt psychology's unanswered questions. In R. J. Sternberg & J. E. Davidson (Eds.), *The nature of insight* (pp. 3–32). Cambridge, MA: MIT Press.
Mayer, R. E. (Ed.). (2005). *The Cambridge handbook of multimedia learning*. New York: Cambridge University Press.
Mayer, R. E. (2008). *Learning and instruction* (2nd ed.). Upper Saddle River, NJ: Pearson Merrill Prentice Hall.
Mayer, R. E. (2009). *Multimedia learning* (2nd ed.). New York: Cambridge University Press.
Mayer, R. E. (2011). *Applying the science of learning*. Upper Saddle River, NJ: Pearson Merrill Prentice Hall.
Mayer, R. E., & DaPra, S.C. (2012). An embodiment effect in computer-based learning with animated pedagogical agents. *Journal of Experimental Psychology: Applied, 18*, 239–252.
Mayer, R. E., & Johnson, C. I. (2010). Adding instructional features that promote learning in a game-like environment. *Journal of Educational Computing Research, 42*, 241–265.
Mayer, R. E., & Wittrock, M. C. (2006). Problem solving. In P. A. Alexander & P. H. Winne (Eds.), *Handbook of educational psychology* (2nd ed., pp. 287–303). Mahwah, NJ: Erlbaum.
O'Neil, H. F. (Ed.). (2005). *What works in distance education: Guidelines*. Greenwich, CT: Information Age Publishing.
Pashler, H., Bain, P., Bottage, B., Graesser, A., Koedinger, K., McDaniel, M., et al. (2007). *Organizing instruction and study to improve student learning (NCER 2007–2004)*. Washington, DC: National Center for Educational Research, Institute of Educational Sciences, U.S. Department of Education. Retrieved from http://ncer.ed.gov

Chapter 9
A Theoretically Based Approach to Cognitive Readiness and Situation Awareness Assessment

Cheryl A. Bolstad, Mica R. Endsley, and Haydee M. Cuevas

Military medical combat teams perform under stressful conditions, constantly dealing with life or death situations, including their own. It is imperative that their leaders have useful decision support tools to determine not only if their team members have the technical knowledge and skills needed to perform successfully but also whether or not they are "cognitively ready" to be deployed. To address this critical need, our research team developed the Medical Cognitive REadiness Survey Tool (M-CREST) to assess the degree to which military medical personnel are cognitively ready to perform their missions effectively. In this chapter, we first begin with a discussion of the theoretical foundation that guided our research project, highlighting the multidimensional nature of the cognitive readiness construct. We then focus more specifically on one important element of cognitive readiness, namely situation awareness (SA). Next, we describe our approach to assessing cognitive readiness (M-CREST) as well as offer recommendations for how to incorporate measures of situation awareness into cognitive readiness assessment. We conclude with implications for future research and development.

C.A. Bolstad, Ph.D. (✉)
SA Technologies, Inc., 76 Lillian Court, Forest Hill, MD 21050, USA
e-mail: cheryl@satechnologies.com

M.R. Endsley
SA Technologies, Inc., 8787 East Lost Gold Circle, Gold Canyon, AZ 85118, USA
e-mail: mica@satechnologies.com

H.M. Cuevas
Embry-Riddle Aeronautical University, 13863 Countryplace Drive, Orlando, FL 32826, USA
e-mail: hmcuevas@earthink.net

H.F. O'Neil et al. (eds.), *Teaching and Measuring Cognitive Readiness*,
DOI 10.1007/978-1-4614-7579-8_9,

9.1 Defining Cognitive Readiness

For our research, it was important to differentiate between military readiness and cognitive readiness. Military readiness is the ability of US military forces to fight and meet the demands of the national military strategy. More specifically, military readiness measures the ability of a unit, such as an Army division or a Navy carrier battle group, to provide the capabilities required by their commanders to successfully execute and achieve their assigned missions (Voith, 2001). This is derived from the ability of each unit to deliver the outputs for which it was designed. Conversely, cognitive readiness is a form of personal mental readiness that supports but does not replace military readiness. Greater levels of cognitive readiness facilitate enhanced cognitive performance, thus enabling military personnel to better perform their assigned duties. In both cases, readiness refers to the "potential" of these individuals and teams to achieve success rather than a measure of their actual success.

Morrison and Fletcher (2002) defined cognitive readiness as the "mental preparation (including skills, knowledge, abilities, motivations, and personal disposition) an individual needs to establish and sustain competent performance in the complex and unpredictable environment of modern military operations" (p. I-3). In our research, we built upon this definition, placing a greater emphasis on individual and team readiness and performance. Conceptually, we define cognitive readiness as possessing the psychological (mental) and sociological (social) knowledge, skills, and attitudes (KSAs) that individuals and team members need to sustain competent professional performance and mental well-being in the dynamic, complex, and unpredictable environments of military operations (Bolstad, Cuevas, Babbitt, Semple, & Vestewig, 2006). While somewhat broad in scope, this definition embraces the KSAs that are needed for superior cognitive readiness while still allowing for the specification of these KSAs to be tailored to targeted team skills and tasks.

9.2 A Multidimensional Construct

Cognitive readiness is a complex, dynamic, multidimensional construct that is formed and maintained when personnel interact with other team members within their operational environment. Therefore, central to understanding cognitive readiness is identifying the factors associated with this construct. Morrison and Fletcher (2002) identified ten factors underlying cognitive readiness: situation awareness, memory, transfer of training, metacognition, automaticity, problem solving, decision making, mental flexibility and creativity, leadership, and emotion. We revised and expanded this list to include several additional factors (see Table 9.1) based on findings from: (1) our critical review of Army combat support hospital after-action reports from Iraq and Afghanistan and the industrial/organizational, human factors,

Table 9.1 Candidate factors related to cognitive readiness drawn from different sources

Individual	Team	Organizational	Environmental
Behavioral style[a]	Cohesion[a]	Management	Fatigue[a]
Cognitive framing[a]	Collective efficacy	Organizational structure	Human–machine interaction
Cognitive resources[a]	Commonality of goals[a]	Satisfaction[a]	Noise
Creativity/flexibility[a]	Communication[a]	Social interactions	Temperature
Decision making[a]	Conflict resolution[a]	Success[a]	Tempo/time pressure
Emotion/anxiety[a]	Coordination	Team size	Uncertainty/ confusion
Experience	Leadership[a]	Training/education[a]	Vibration
Global response to stress[a]	Shared cognition		Work–rest cycles
Individual roles[a]	Shared mental models		Workload[a]
Memory capacity[a]	Team social processes[a]		
Mental models			
Metacognition			
Problem solving abilities[a]			
Self-efficacy			
Situation awareness			

[a]Factors rated by military medical personnel

management, and team performance literature; and (2) our interviews with military medical personnel (surgeons, medics, and nurses) about their experiences working in teams and what factors they found affected team performance and cognitive readiness (for a detailed description of this research, see Bolstad, Babbitt, & Semple, 2004). We were specifically interested in the identification of factors that best predict good team performance during dynamic and stressful situations, such as military combat operations. Based on this need, we paired down the list to a smaller subset of 21 factors (see factors marked with an asterisk in Table 9.1). We then asked 12 military medical staff members, who were currently deployed to Afghanistan and working in the same combat support hospital, to rate these 21 factors in terms of importance in performing their work (for a detailed description of this study, see Bolstad et al., 2004, 2006). Participants rated all 21 factors as moderately to highly important. None were rated as "unimportant" to pre-deployment cognitive readiness.

Building on this earlier work, we developed a theoretical framework that illustrates how many of these factors interact with other individual, team, organizational, and environmental factors to influence cognitive readiness (see Fig. 9.1). Several of these factors are inherent in the individual team member, such as, for example, memory capacity and other cognitive resources, problem solving and decision-making ability, and creativity and flexibility. Others are more relevant at the team level, such as communication, shared mental models, commonality of goals, and leadership. Although all these factors are important to cognitive readiness, in this chapter, we will focus more specifically on *situation awareness*, discussed next.

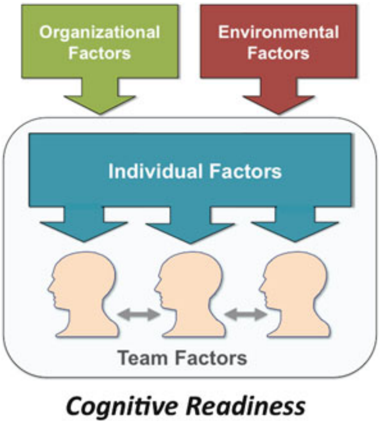


Fig. 9.1 Theoretical framework illustrating categories of factors influencing cognitive readiness

9.3 Situation Awareness and Cognitive Readiness

Situation awareness (SA) involves being aware of what is happening around you to understand how information, events, and your own actions will affect your goals and objectives, both now and in the near future. More formally, SA can be defined as "the perception of elements in the environment within a volume of time and space, the comprehension of their meaning, and the projection of their status in the near future" (Endsley, 1995b, p. 36). Having complete, accurate, and up-to-the-minute SA is considered to be essential in any domain where the effects of ever increasing technological and situational complexity on the human decision-maker are a concern (Endsley, 1995b). SA has been recognized as a critical, yet often elusive, foundation for successful decision making across a broad range of complex and dynamic systems, including aviation and air traffic control (e.g., Endsley, 2009; Nullmeyer, Stella, Montijo, & Harden, 2005), emergency response and military command and control operations (e.g., Blandford & Wong, 2004; Gorman, Cooke, & Winner, 2006), railroad operations (e.g., Golightly, Wilson, Lowe, & Sharples, 2010; Roth, Multer, & Raslear, 2006); power transmission and distribution (Connors, Endsley, & Jones, 2007; Salmon et al., 2008); and offshore oil and nuclear power plant management (e.g., Flin & O'Connor, 2001). Indeed, lacking SA or having inadequate SA has been consistently identified as one of the primary factors in accidents attributed to human error (e.g., Hartel, Smith, & Prince, 1991; Merket, Bergondy, & Cuevas-Mesa, 1997; Nullmeyer et al., 2005; Stanton, Chambers, & Piggott, 2001). Thus, SA is especially crucial in domains where information flow can be quite high and poor decisions may lead to serious consequences (e.g., piloting an airplane, functioning as a soldier, or treating critically ill or injured patients). Accordingly, it is not surprising that SA is an important component of cognitive readiness.

As described earlier, cognitive readiness represents the "potential" of team members to be ready to perform their job and SA is an essential element of this readiness. However, within the context of cognitive readiness, the focus is not on

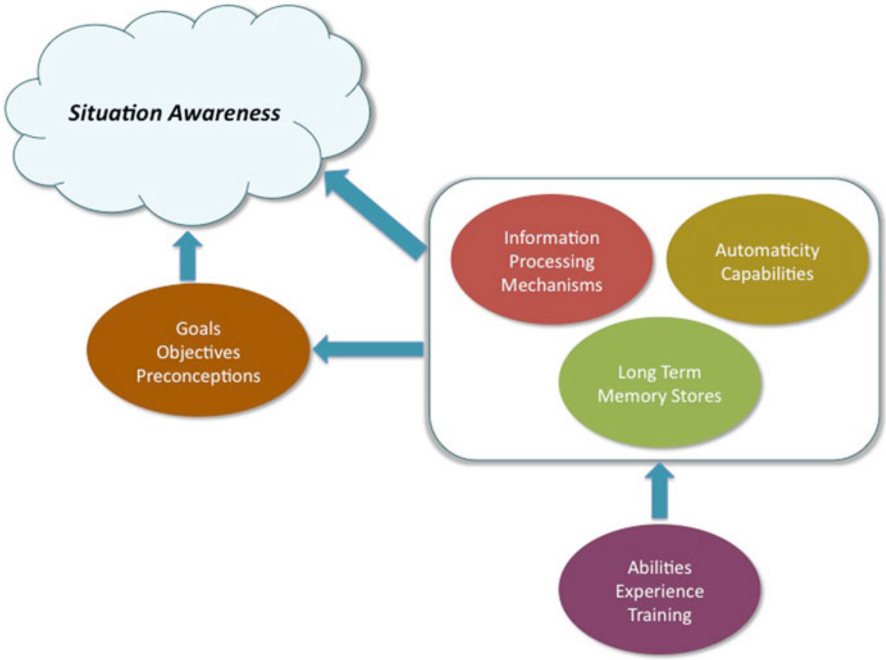


Fig. 9.2 Individual factors that affect situation awareness according to Endsley's (1995b) theoretical model

actual SA, but rather on the "potential" of an individual or team of individuals to achieve SA. The question then becomes what makes some individuals more likely to have better SA than others? Endsley (1995b) constructed a theoretical model of SA to describe how SA is formed and how it affects decision making and performance. As shown in Fig. 9.2, several individual factors influence SA. Many of these same factors also influence cognitive readiness (refer to Fig. 9.1).

Research has shown evidence for specific individual differences in abilities that potentially affect SA development. For example, Gugerty, Brooks, and Treadaway (2004) demonstrated how individual differences in perceptual and cognitive abilities (e.g., mental rotation, working memory capacity, divided and selective attention) are related to the performance of specific subtasks (e.g., navigation, maneuvering) in transportation operations (flying and driving) and to the ability to maintain SA while performing transportation tasks (see also Sohn & Doane, 2004). Similarly, other studies have shown that individuals who are better able to share attention on tasks exhibit better SA (Endsley & Bolstad, 1994; Gugerty & Tirre, 1997). In addition, Endsley and Bolstad (1994) found that individuals who performed better on psychomotor tasks also demonstrated better SA. While psychomotor skills may not be directly related to the cognitive task of developing and maintaining SA, it is hypothesized that greater levels of skill and automaticity in psychomotor tracking may free up the cognitive resources needed for SA.

Knowledge and information stored in long-term memory can also affect an individual's potential to form SA. Specifically, to understand patterns or trends in elements perceived in the environment, individuals retrieve analogous instances from their long-term memory that can then be used to compare against the current situation (Serfaty, MacMillan, Entin, & Entin, 1997; Sohn & Doane, 2004). Similarly, these long-term memory stores are also used to support predictions of future trends or changes in the situation. This repertoire of conceptual patterns or "mental models" stored in long-term memory is expanded through the development of expertise (Feltovich, Prietula, & Ericsson, 2006; Glaser, 1989; Smith, Ford, & Kozlowski, 1997). Thus, not surprisingly, expertise is related to better SA (Endsley, 2006). Experts, as compared to novices, are likely to have developed a greater sensitivity for detecting or recognizing patterns in specific types of data through training, extensive experience, better focused attention, or more effective use of data representations (Endsley, 1997; Garrett & Caldwell, 2009). In turn, this greater sensitivity may enable them to detect events with more accuracy and speed.

Beyond abilities, an individual's goals and objectives also directly influence the development of SA. Specifically, an individual's goals have a bearing on which specific cues in the environment are perceived, that is, individuals selectively direct their attention to information that is relevant to their goals and tend to ignore environmental cues that may not be as pertinent. Similarly, critical for developing and maintaining higher levels of SA is understanding how the current and future state of the situation affects one's goals.

To garner a better understanding of cognitive readiness and its constituent factors, valid and reliable measures of these constructs are absolutely indispensable. Although cognitive readiness and SA are both influenced by several similar factors, these constructs are also affected by different individual, task, and environmental factors. Thus, SA measures cannot be used to generally assess cognitive readiness and vice versa. Numerous well established measures of SA exist (see Endsley & Garland, 2000; Fracker, 1991a, 1991b; Salmon, Stanton, Walker, & Green, 2006; Wright, Taekman, & Endsley, 2004), yet very little research has been conducted on assessing cognitive readiness. Our research was aimed at addressing this important issue. Next, we describe our approach to assessing cognitive readiness as well as briefly discuss different approaches to assessing SA.

9.4 Assessing Cognitive Readiness

One approach to assessing cognitive readiness is to focus on members' subjective evaluation of their team's operational readiness. For example, Guerlain et al. (2004) developed a self-evaluation questionnaire that asked team members to rate their team's readiness to perform a mission or activity in terms of a suitable plan, sufficiency of personnel/skill sets, effective leadership, and effective communication. While this approach may have face validity, it cannot always ensure predictive validity. In contrast, our assessment approach is based on a theoretical framework that highlights the multidimensional nature of the cognitive readiness construct. In this section, we briefly describe the design, development, and initial usability assessment of our M-CREST (for a more detailed description, see Bolstad, Cuevas, Costello, & Babbitt, 2008).

Our theoretically based design approach involved developing a prototype measurement system that provides valid assessment of several essential factors that are indicative of successful performance at both the individual and team level. We focused on the more stable, enduring factors that are internal to the operator, that is, individual and team competencies (or KSAs), rather than organizational or environmental factors (cf. Cannon-Bowers, Tannenbaum, Salas, & Volpe, 1995). Although vital for optimizing performance, certain organizational (e.g., organizational structure, training/educational opportunities) and environmental (e.g., noise, temperature, task load) constraints are beyond the control of the operator and change from moment to moment, team to team, and mission to mission. Thus, it is doubtful that one could reliably assess performance under all these varying conditions.

The M-CREST prototype was designed to assess 12 essential factors associated with cognitive readiness. To facilitate the selection of the appropriate measures, we organized these 12 factors within categories that reflect how they are related to one another in terms of the individual's KSAs and team orientation (see Table 9.2). We then conducted an exhaustive search for off-the-shelf, validated measurement instruments from the education, business, and psychology domains (for a detailed description, see Bolstad et al., 2007). We developed and applied a comprehensive set of evaluation criteria to narrow down this list to six instruments to include in the design of our initial prototype. The M-CREST prototype surveys team members on 12 essential cognitive readiness factors, automatically scores their responses in real time, and provides recommendations to enhance their cognitive readiness.

The M-CREST interface consists of three components: user, administrator, and report generator. The entire system resides on a physically and electronically secure server and uses industry-standard software. All three interfaces are remotely accessible from virtually any location provided the user has an Internet connection and a standard Internet browser. Individual survey takers interact with M-CREST via the *user interface*, responding to each survey item using Likert-type scales. Upon completion of all the survey items, individuals receive their personalized M-CREST Individual Report (see Fig. 9.3), which provides the following information for each cognitive readiness factor surveyed:

- Overall score in terms of high, moderate, or low
- Layman's definition of the factor
- Description of the factor's relevance
- List of key KSAs that help ensure cognitive readiness for this area
- List of individual and team benefits associated with high readiness on this factor
- Up to three URLs for web sites presenting content intended to help individuals further enhance their cognitive readiness with respect to this factor

Table 9.2 Twelve cognitive readiness factors assessed by M-CREST prototype

Individual KSAs	Definition
Knowledge	
Thinking and planning strategies	How one approaches solving a problem, including the process of defining the problem to be solved, identifying the requirements (what information and actions are needed) for solving the problem, and effectively applying the appropriate techniques or strategies with the objective of solving the problem (O'Neil & Abedi, 1996)
Monitoring and self-checking strategies	Conscious and periodic self-checking of whether one's goal is being achieved, and, when necessary, selecting and applying different strategies (O'Neil & Abedi, 1996). Often referred to as metacognition or metacognitive skills
Skills	
Leadership	Ability to positively influence group members so as to help achieve the goals of the group (Kauses & Posner, 2001)
Individual roles on the job	Accepted, mandated, or assigned behaviors associated with a particular position within a group (Cammann, Fichman, Jenkins, & Klesh, 1983)
Attitudes	
Behavioral style	Attitude- and personality-driven patterns of behavior people exhibit in work and social settings (Cornelius, 2009)
Dealing with stress	How one manages stress in general, especially in ambiguous situations or when one must solve difficult problems (Heppner, 1988)
Flexibility/openness	Ability to be open to ideas that are different from one's own and to people who are different from oneself (Kelley & Meyers, 1995)
Self-confidence	People's judgments of their capabilities to organize and execute courses of actions required to attain designated types of performances … a judgment of one's capability to accomplish a certain level of performance (Bandura, 1986). More commonly referred to as self-efficacy
Team orientation	
Team cohesion	Active involvement and commitment driving the willingness to remain, and freely interact, in a group (Mullen & Copper, 1994)
Team common goals	Degree to which specific individual, team, or organizational goals are shared by members of a group (Stevens & Campion, 1994)
Team confidence	Members' shared belief in their team's ability or competence to perform a task or attain desired outcomes (Bandura, 1986; Pethe, 2002). More commonly referred to as collective efficacy
Team cooperation	Willingness on the part of team members to engage in coordinative or adaptive behavior; represents the attitudinal component underlying team coordination (Fiore, Salas, Cuevas, & Bowers, 2003)

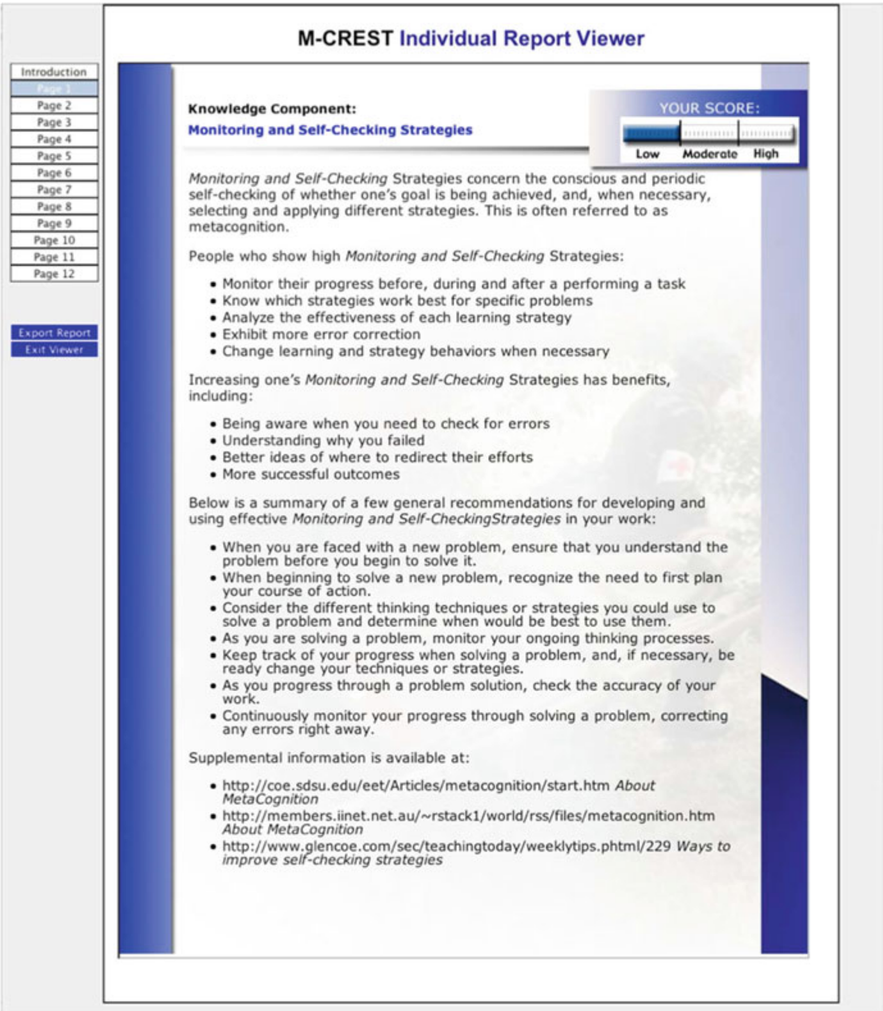


Fig. 9.3 Example page of M-CREST Individual Report

Supervisors interact with the *administrator interface* to identify and organize survey takers into groups and select the targeted cognitive readiness factors to be surveyed for each group. Once all group members have completed the survey, supervisors then access the *report generator interface* to view group-level summaries of the M-CREST survey results. Supervisors can request reports that compare a particular group's results on some or all the KSAs surveyed or compare different groups on specific KSA (see Fig. 9.4). It should be noted that the content of the M-CREST survey items and reports were not specific to military medical teams, but rather were domain-general, that is, were written using general terminology applicable to any domain.

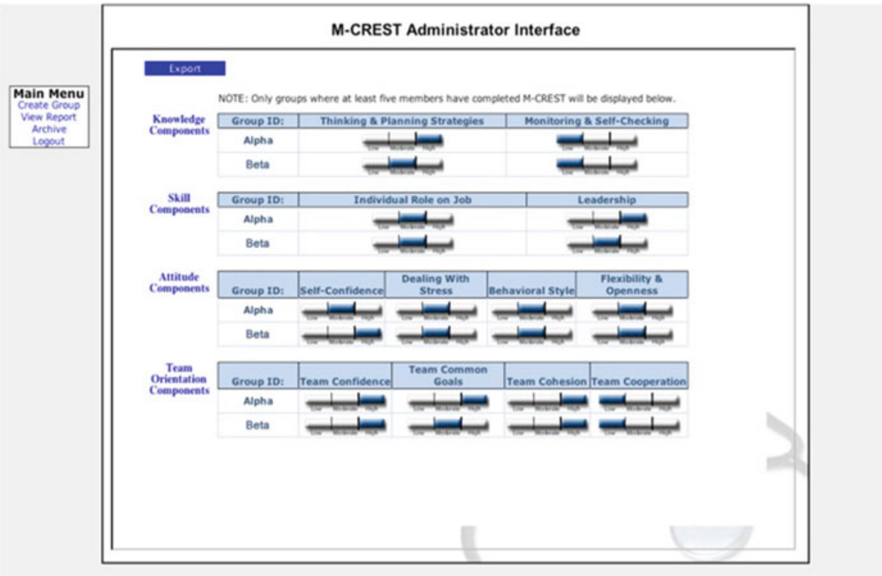


Fig. 9.4 Example of the M-CREST report generator output

The initial usability assessment of our M-CREST prototype was an important part of the development process, as feedback from potential users is critical to ensure that we meet our objective of designing a tool that can be used to assess and enhance cognitive readiness (for a detailed description of this study, see Bolstad et al., 2008). We did not have access to military medical teams for the usability evaluation of the M-CREST prototype. Thus, to increase the generalizability of the study's findings to military medical personnel, we solicited participation from an operationally relevant civilian population. Participants were recruited from three different test sites, and included firefighters, emergency medical technicians, and first responders (although nonmilitary, these individuals also perform in complex, high risk environments). Seven participants (all males; average age 45.4 years) from the three test sites participated in this study. Participants reported an average of 9.2 years on their current job and an average of 15.8 years in their career field. Participants were asked to complete nine sets of questions (total of 117 questions) hosted on the M-CREST prototype and were then presented with an M-CREST Individual Report summarizing their responses to the M-CREST survey as well as providing recommendations to enhance their cognitive readiness. Participants were also asked to provide feedback on their interaction with the M-CREST software, including ease of use and usefulness of information provided in their M-CREST Individual Report. Overall, participants found it easy to understand the M-CREST survey items, found the information and recommendations on enhancing their cognitive readiness presented in their M-CREST Individual Report relevant to their work, and were satisfied with the information they gained by participating in the study. Team leaders reported finding the interface easy to use and the information helpful for better understanding their teams.

9.5 Assessing Situation Awareness

Encouraged by the promising results of our initial usability assessment, we set out to explore the feasibility of expanding M-CREST to include measures of other essential cognitive readiness factors, such as SA. In this section, we offer recommendations for how to incorporate measures of SA into cognitive readiness assessment. In general, methodologies to assess SA vary in terms of direct measurement (e.g., objective real-time probes or subjective questionnaires assessing perceived SA) or indirect methods (e.g., process indices, trained observer ratings) that infer SA based on operator physiological states, behavior, or performance. Direct measures are typically considered to be "product-oriented" in that these techniques assess an SA outcome; indirect measures are considered to be "process-oriented," focusing on the underlying processes or mechanisms required to achieve SA (Graham & Matthews, 2000). The selection of which methodology to use depends upon the researcher's objectives and what data collection facilities or setup are available.

For example, the Situation Awareness Global Assessment Technique (SAGAT; Endsley, 1995a) provides direct objective measurement of SA by comparing an individual's perceptions of the situation or environment to some "ground truth" reality. SAGAT involves temporarily halting a simulation or operational activity at randomly selected times and removing task information sources (e.g., blacking out information displays); administering a set of queries that target each individual's dynamic SA information requirements (i.e., what they need to know at that point in time) with respect to the domain of interest; and resuming the simulation or activity (Endsley, 1995a). For settings in which disruptions to task performance are not practical or desirable, real-time probes (e.g., open-ended questions embedded as verbal communications during the task) can be administered to naturally and unobtrusively assess operator SA (Jones & Endsley, 2000). Real-time probes are similar to SAGAT in that these query operators on their knowledge of key task-relevant information in the environment; however, this methodology differs from the SAGAT in that task performance is not disrupted (i.e., the simulation or task is not stopped) but rather the queries are incorporated as a natural part of the task. Process-oriented indirect measures, such as the Situation Awareness Behaviorally Anchored Rating Scale (SABARS; Strater, Endsley, Pleban, & Matthews, 2001), also do not require interrupting task performance. Instead, these measures involve unobtrusive ratings by expert-trained observers of the types of overt team behaviors and communications that are indicative of good SA.

Because M-CREST is designed to be completed by individuals removed from the operational environment (i.e., not during task performance), assessment of dynamic constructs such as SA instead must rely on proxy measures to evaluate operators' predicted response in a given hypothetical situation. For example, short

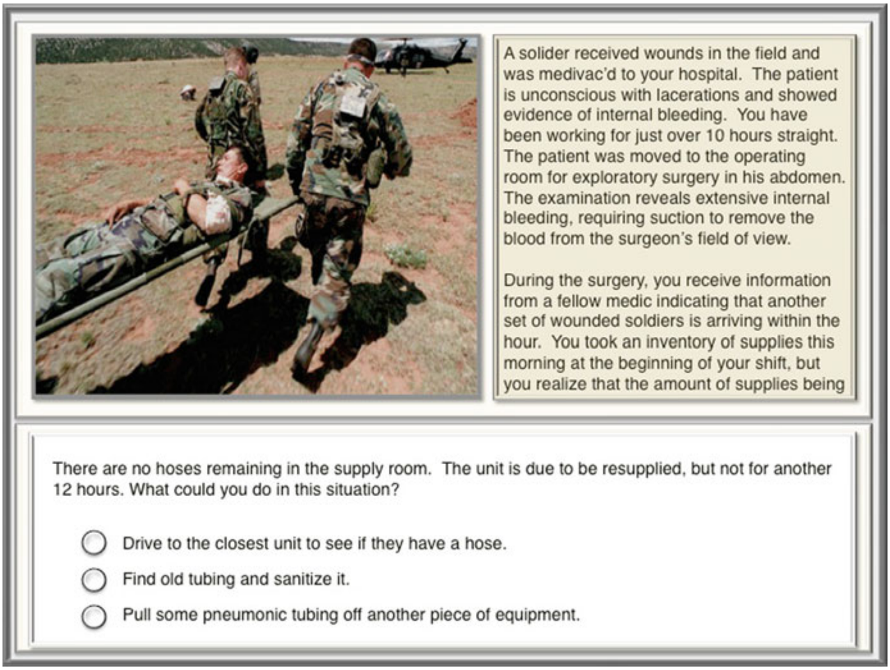


Fig. 9.5 Illustrative example of a scenario-based survey item for assessing situation awareness using M-CREST

text-based vignettes can be presented on the M-CREST that ask operators to predict what they would do or how they would react in a given situation (see Fig. 9.5). Responses for each individual and collectively for the team can then be examined to see if they have a shared understanding of the situation and know how best to respond. This could be used to gauge the "potential" of team members to develop SA when performing in the operational environment. The validity and utility of this scenario-based approach to assessment using text-based vignettes has been well documented (cf. Cannon-Bowers, Burns, Salas, & Pruitt, 1998; Rosen et al., 2008; Salas & Cannon-Bowers, 2000; Schmorrow, Cohn, & Nicholson, 2009; Vincenzi, Wise, Mouloua, & Hancock, 2008).

However, creating text-based vignettes for predictive SA assessment requires that scenarios be sufficiently detailed to engage the individual. In addition, to be operationally relevant to team performance, the vignettes must be specific to the team's domain. As such, developing metrics for this type of assessment can be very time-consuming and labor-intensive and the scenarios may not be readily generalizable to other domains. While M-CREST currently utilizes only domain-general measures, we, nevertheless, realize that incorporating domain-specific measures of cognitive readiness factors, such as SA, is absolutely vital to improve our tool's utility for providing a more comprehensive assessment of cognitive readiness.

9.6 Future Directions

Cognitive readiness is a new and continually evolving construct. Many questions still remained unanswered regarding what factors determine cognitive readiness and how cognitive readiness influences actual performance in the operational environment. Addressing these questions will require identifying existing as well as developing new valid and reliable measures of cognitive readiness and its constituent factors, which can then be incorporated into the design of M-CREST. Future research, therefore, is clearly warranted to establish the construct and convergent validity of measures for this important construct as well as the predictive validity of these measures with regard to individual and team performance. For example, operational assessments can be conducted with military medical personnel from Army combat support hospitals completing training at an Army Trauma Training Center as well as nonmilitary medical residents working at a high volume emergency room in a civilian hospital to statistically evaluate the psychometric properties of the survey items included in a more fully developed M-CREST (e.g., conduct a factor analysis of survey responses). Follow-up studies with participants can also be performed to determine if the feedback provided in the M-CREST Individual Report was utilized and proved helpful on the job.

M-CREST has been designed to enhance a team's cognitive readiness by drawing their attention to important KSAs that will enable them to more effectively deal with their new environment and responsibilities as well as improve their interactions with their team members and others in the field. For example, our research with military medical teams revealed several KSAs essential to team performance including problem solving, decision making, situation awareness, leadership, communication, and team cohesion as well as highlighted the importance of considering the effects of other factors such as fatigue, workload, and stress. M-CREST has also been designed to provide useful decision support to team leaders by identifying their team's strengths and weaknesses with regard to their cognitive readiness prior to task performance. As such, M-CREST can also potentially be used to balance or create teams based on their scores on different cognitive readiness factors. Further, because of its modular design, the surveys administered via M-CREST can be flexibly tailored to assess the KSAs deemed most critical for a given team or operational domain. Following assessment, M-CREST can then be used to evaluate the effectiveness of training interventions targeted at improving these specific areas of a team's cognitive readiness. Coupled with validated training programs, M-CREST, therefore, represents a valuable decision support tool that team leaders can use to prepare their teams to ensure successful performance in the operational environment.

Although M-CREST was originally designed for military medical training organizations and medical teams deploying worldwide, cognitive readiness is also applicable to Homeland Security, law enforcement, emergency, first responders, and other civilian medical personnel. Indeed, our initial usability assessment with firefighters, emergency medical technicians, and first responders demonstrated both M-CREST's potential usefulness for this population and its utility as a domain-general assessment tool. Thus, our development plans entail enhancing the design of M-CREST to make it applicable to a wider population by incorporating measures of other essential cognitive readiness factors, such as situation awareness. We also changed the product name to T-CREST (Team Cognitive REadiness Survey Tool) to reflect the nonmedical language contained in the survey items.

9.7 Conclusions

Human performance in today's technologically complex operations is influenced by a broad range of individual, team, organizational, and environmental factors. Therefore, from an applied perspective, cognitive readiness focuses on defining and optimizing the human dimension of the sociotechnical system by ensuring that individuals and teams possess the essential KSAs needed to perform effectively in these challenging domains (cf. Bowman & Thomas, 2008). Of particular interest is a team's cognitive readiness to maintain performance in foreign cultures, adverse climates, and demanding uncertain circumstances. Psychological researchers will play a vital role in helping optimize human performance through an understanding of the cognitive, behavioral, and attitudinal factors underlying individual and team performance, and through the identification of valid measures to assess these essential KSAs. The line of research reported in this chapter represents a theoretically based, operationally valid approach to addressing this important objective.

Acknowledgements Work on the research reported in this chapter was partially supported by a Phase I and Phase II Small Business Innovative Research Contract (W81XWH-04-C-0014) awarded to the first author from the Office of the Secretary of Defense (OSD). The views and conclusions contained herein, however, are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the OSD, U.S. Army, Department of Defense, U.S. Government, or the organizations with which the authors are affiliated. We wish to express special thanks to Anthony M. Costello, Bettina Babbitt, Clarence Semple, and Richard Vestewig for their vital contributions to this project. Portions of this chapter were presented at: the International Ergonomics Association 16th World Congress on Ergonomics, Maastricht, The Netherlands; the Human Factors and Ergonomics Society 52nd Annual Meeting, New York, NY; and the 2009 Teaching and Measuring Cognitive Readiness Workshop, Los Angeles, CA, which was partially supported by a grant from the Office of Naval Research (ONR), Award Number N000140810126; however, the findings and opinions expressed here do not necessarily reflect the positions or policies of the ONR.

References

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice-Hall.

Blandford, A., & Wong, W. (2004). Situation awareness in emergency medical dispatch. *International Journal of Human Computer Studies, 61*, 421–452.

Bolstad, C. A., Babbitt, B. A., & Semple, C. A. (2004). *Predicting cognitive performance of deploying health teams* (Final Report prepared for the U.S. Army Medical Research and Material Command, under Contract No. W91XWH-C-004). Marietta, GA: SA Technologies.

Bolstad, C. A., Cuevas, H. M., Babbitt, B. A., Semple, C. A., & Vestewig, R. E. (2006). Predicting cognitive readiness of military health teams. In *Proceedings of the International Ergonomics Association 16th World Congress on Ergonomics [CD-ROM]*, Maastricht, The Netherlands, July 10–14, 2006. Amsterdam: Elsevier.

Bolstad, C. A., Cuevas, H. M., Costello, A., Babbitt, B. A., Semple, C. A., & Vestewig, R. E. (2007). *Final report for SBIR Phase II: Predicting cognitive performance of deploying health teams* (Technical Report for the Office of the Secretary of Defense (OSD), prepared under Contract No. W81XWH-04-C-0014). Marietta, GA: SA Technologies.

Bolstad, C. A., Cuevas, H. M., Costello, A. M., & Babbitt, B. A. (2008). Predicting cognitive readiness of deploying military medical teams. In *Proceedings of the Human Factors and Ergonomics Society 52nd Annual Meeting* (pp. 970–974). Santa Monica, CA: Human Factors and Ergonomics Society.

Bowman, E. K., & Thomas, J. A. (2008). C2 of unmanned systems in distributed ISR operations. In *Proceedings of the 13th International Command and Control Research and Technology Symposia* (ICCRTS 2008), June 17–19, 2008, Seattle, WA.

Cammann, C., Fichman, M., Jenkins, G. D., & Klesh, J. R. (1983). Assessing the attitudes and perceptions of organization members. In S. E. Seashore, E. E. Lawler III, P. H. Mirvis, & C. Cammann (Eds.), *Assessing organizational change: A guide to methods, measures, and practices* (pp. 71–138). New York: Wiley.

Cannon-Bowers, J. A., Burns, J. J., Salas, E., & Pruitt, J. S. (1998). Advanced technology in scenario based training. In J. A. Cannon-Bowers & E. Salas (Eds.), *Making decisions under stress: Implications for individual and team training* (pp. 365–374). Washington, DC: American Psychological Association.

Cannon-Bowers, J. A., Tannenbaum, S. I., Salas, E., & Volpe, C. E. (1995). Defining competencies and establishing team training requirements. In R. A. Guzzo & E. Salas (Eds.), *Team effectiveness and decision making in organizations* (pp. 333–380). San Francisco, CA: Jossey-Bass.

Connors, E. S., Endsley, M. R., & Jones, L. (2007). Situation awareness in the power transmission and distribution industry. In *Proceedings of the Human Factors and Ergonomics Society 51st Annual Meeting* (pp. 215–219). Santa Monica, CA: Human Factors and Ergonomics Society.

Cornelius, E. T. (2009). *Leadership styles for dealing with people. Part 1: Identifying your personal style in dealing with others*. Cornelius and Associates. Retrieved March 18, 2011, from http://www.corneliusassoc.com/library.html

Endsley, M. R. (1995a). Measurement of situation awareness in dynamic systems. *Human Factors, 37*(1), 65–84.

Endsley, M. R. (1995b). Toward a theory of situation awareness in dynamic systems. *Human Factors, 37*(1), 32–64.

Endsley, M. R. (1997). The role of situation awareness in naturalistic decision making. In C. E. Zsambok & G. Klein (Eds.), *Naturalistic decision making* (pp. 269–283). Mahwah, NJ: Lawrence Erlbaum.

Endsley, M. R. (2006). Expertise and situation awareness. In K. A. Ericsson, N. Charness, R. R. Hoffman, & P. J. Feltovich (Eds.), *The Cambridge handbook of expertise and expert performance* (pp. 633–651). New York, NY: Cambridge University Press.

Endsley, M. R. (2009). Situation awareness in aviation. In J. A. Wise, V. D. Hopkin, & D. J. Garland (Eds.), *Handbook of aviation human factors* (2nd ed., pp. 12-1–12-21). Boca Raton, FL: CRC Press—Taylor & Francis.

Endsley, M. R., & Bolstad, C. A. (1994). Individual differences in pilot situation awareness. *The International Journal of Aviation Psychology, 4*(3), 241–264.

Endsley, M. R., & Garland, D. J. (Eds.). (2000). *Situation awareness analysis and measurement*. Mahwah, NJ: Lawrence Erlbaum Associates.

Feltovich, P. J., Prietula, M. J., & Ericsson, K. A. (2006). Studies of expertise from psychological perspectives. In K. A. Ericsson, N. Charness, R. R. Hoffman, & P. J. Feltovich (Eds.), *The Cambridge handbook of expertise and expert performance* (pp. 41–67). New York, NY: Cambridge University Press.

Fiore, S. M., Salas, E., Cuevas, H. M., & Bowers, C. A. (2003). Distributed coordination space: Toward a theory of distributed team process and performance. *Theoretical Issues in Ergonomics Science, 4*, 340–363.

Flin, R., & O'Connor, P. (2001). Applying crew resource management in offshore oil platforms. In E. Salas, C. A. Bowers, & E. Edens (Eds.), *Improving teamwork in organization: Applications of resource management training* (pp. 217–233). Mahwah, NJ: Erlbaum.

Fracker, M. L. (1991a). *Measures of situation awareness: An experimental evaluation* (Report No. AL-TR-1991-0127). Wright-Patterson Air Force Base, OH: Armstrong Laboratories.

Fracker, M. L. (1991b). *Measures of situation awareness: Review and future directions* (Report No. AL-TR-1991-0128). Wright-Patterson Air Force Base, OH: Armstrong Laboratories.

Garrett, S. K., & Caldwell, B. S. (2009). Human factors aspects of planning and response to pandemic events. In *Proceedings of the 2009 Industrial Engineering Research Conference*, May 30–June 3, 2009, Miami, FL.

Glaser, R. (1989). Expertise and learning: How do we think about instructional processes now that we have discovered knowledge structures? In D. Klahr & K. Kotovsky (Eds.), *Complex information processing: The impact of Herbert A. Simon* (pp. 269–282). Hillsdale, NJ: Erlbaum.

Golightly, D., Wilson, J. R., Lowe, E., & Sharples, S. (2010). The role of situation awareness for understanding signalling and control in rail operations. *Theoretical Issues in Ergonomics Science, 11*(1), 84–98.

Gorman, J. C., Cooke, N. J., & Winner, J. L. (2006). Measuring team situation awareness in decentralized command and control environments. *Ergonomics, 49*(12–13), 1312–1325.

Graham, S. E., & Matthews, M. D. (2000). Modeling and measuring situation awareness. In J. H. Hiller & R. L. Wampler (Eds.), *Workshop on assessing and measuring training performance effectiveness* (Tech. Rep. 1116) (pp. 14–24). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Guerlain, S., O'Hargan, K., Lander, E., Gibson, W., Chang, J., Donohue, S., et al. (2004). Predicting the cognitive readiness of Army search and rescue teams. In *Proceedings of the International Conference on Systems, Man and Cybernetics (IEEE SMC 2004)*, October 10–13, 2004, The Hague, Netherlands.

Gugerty, L., Brooks, J. O., & Treadaway, C. A. (2004). Individual differences in situation awareness for transportation tasks. In S. Banbury & S. Tremblay (Eds.), *A cognitive approach to situation awareness: Theory and application* (pp. 193–212). Aldershot, England: Ashgate.

Gugerty, L., & Tirre, W. (1997). Situation awareness: A validation study and investigation of individual differences. In *Proceedings of the Human Factors and Ergonomics Society 40th Annual Meeting* (pp. 544–568). Santa Monica, CA: Human Factors and Ergonomics Society.

Hartel, C. E. J., Smith, K., & Prince, C. (1991, April). *Defining aircrew coordination: Searching mishaps for meaning*. Paper presented at the 6th International Symposium on Aviation Psychology, Columbus, OH.

Heppner, P. P. (1988). *The Problem Solving Inventory (PSI): Manual*. Palo Alto, CA: Consulting Psychologists Press.

Jones, D. G., & Endsley, M. R. (2000). Examining the validity of real-time probes as a metric of situation awareness. In *Proceedings of the International Ergonomics Association 14th World Congress on Ergonomics and the Human Factors and Ergonomics Society 44th Annual Meeting*. Santa Monica, CA: Human Factors and Ergonomics Society.

Kauses, J. M., & Posner, B. Z. (2001). *Leadership practices inventory*. San Francisco, CA: Jossey-Bass/Pfeiffer.

Kelley, C., & Meyers, J. (1995). *The Cross-Cultural Adaptability Inventory (manual)*. Minneapolis, MN: National Computer Systems.

Merket, D. C., Bergondy, M., & Cuevas-Mesa, H. (1997, March). *Making sense out of teamwork errors in complex environments.* Paper presented at the 18th Annual Industrial/Organizational-Organizational Behavior Graduate Student Conference, Roanoke, VA.

Morrison, J. E., & Fletcher, J. D. (2002). *Cognitive readiness* (IDA Paper P-3735). Alexandria, VA: Institute for Defense Analyses.

Mullen, B., & Copper, C. (1994). The relation between group cohesiveness and performance: An integration. *Psychological Bulletin, 115*, 210–227.

Nullmeyer, R. T., Stella, D., Montijo, G. A., & Harden, S. W. (2005). Human factors in Air Force flight mishaps: Implications for change. In *Proceedings of the 27th Annual Interservice/Industry Training, Simulation, and Education Conference* (paper no. 2260). Arlington, VA: National Training Systems Association.

O'Neil, H. F., Jr., & Abedi, J. (1996). Reliability and validity of a State Metacognitive Inventory: Potential for alternative assessment. *The Journal of Educational Research, 89*(4), 234–245.

Pethe, S. (2002). Collective efficacy: Development of a measure. *Decision, 29*(1), 25–35.

Rosen, M. A., Salas, E., Wilson, K. A., King, H. B., Salisbury, M., Augenstein, J. S., et al. (2008). Measuring team performance in simulation-based training: Adopting best practices for healthcare. *Simulation in Healthcare, 3*(1), 33–41.

Roth, E. M., Multer, J., & Raslear, T. (2006). Shared situation awareness as a contributor to high reliability performance in railroad operations. *Organization Studies, 27*(7), 967–987.

Salas, E., & Cannon-Bowers, J. A. (2000). The anatomy of team training. In S. Tobias & J. D. Fletcher (Eds.), *Training and retraining: A handbook for business, industry, government, and the military* (pp. 312–335). New York, NY: Macmillan.

Salmon, P., Stanton, N., Walker, G., & Green, D. (2006). Situation awareness measurement: A review of applicability for C4I environments. *Applied Ergonomics, 37*(2), 225–238.

Salmon, P. M., Stanton, N. A., Walker, G. H., Jenkins, D. P., Baber, C., & McMaster, R. (2008). Representing situation awareness in collaborative systems: A case study in the energy distribution domain. *Ergonomics, 51*(3), 367–384.

Serfaty D., MacMillan J., Entin, E. E., & Entin E. B. (1997). The decision-making expertise of battle commanders. In C. E. Zsambok and G. Klein (Eds.), *Naturalistic decision-making* (pp. 233–246). New York: Lawrence Erlbaum.

Schmorrow, D., Cohn, J., & Nicholson, D. (Eds.). (2009). *The PSI handbook of virtual environments for training and education: Developments for the military and beyond*. Westport, CT: Praeger Security International.

Smith, E. M., Ford, J. K., & Kozlowski, S. W. J. (1997). Building adaptive expertise: Implications for training design strategies. In M. A. Quinones & A. Ehrenstein (Eds.), *Training for a rapidly changing workplace: Applications of psychological research* (pp. 89–118). Washington, DC: American Psychological Association.

Sohn, Y. W., & Doane, S. M. (2004). Memory processes of flight situation awareness: Interactive roles of working memory capacity, long-term working memory, and expertise. *Human Factors, 46*(3), 461–475.

Stanton, N. A., Chambers, P. R. G., & Piggott, J. (2001). Situational awareness and safety. *Safety Science, 39*, 189–204.

Stevens, M. J., & Campion, M. A. (1994). The knowledge, skill, and ability requirements for teamwork: Implications for human resource management. *Journal of Management, 20*(2), 503–530.

Strater, L. D., Endsley, M. R., Pleban, R. J., & Matthews, M. D. (2001). *Measures of platoon leader situation awareness in virtual decision making exercises* (Research Report No. 1770). Alexandria, VA: Army Research Institute.

Vincenzi, D. A., Wise, J. A., Mouloua, M., & Hancock, P. A. (Eds.). (2008). *Human factors in simulation and training*. Boca Raton, FL: CRC Press.

Voith, M. R. (2001). Military readiness. *The Army Doctrine and Training Bulletin, 4*(2), 41–48.

Wright, M. C., Taekman, J. M., & Endsley, M. R. (2004). Objective measures of situation awareness in a simulated medical environment. *Quality & Safety in Health Care, 13*(Suppl 1), i65–i71.

Chapter 10
Adaptive Expertise and Cognitive Readiness: A Perspective from the Expert-Performance Approach

K. Anders Ericsson

The interest in creativity, general problem solving, and general intelligence has remained relatively stable across the last few centuries of research in psychology and psychological science. On several occasions investigators have proposed that general problem-solving and decision-making abilities might be trainable and the research interest increased. One of these proposals was made by the gestalt psychologist, Wertheimer (1945/1982) who distinguished "learning by drill, by external association, by external conditioning, by memorizing, and by trial and error" from "developing structural insight, structural mastery, and meaningful learning in the real sense of the word" (p. 246).

Unfortunately, researchers, who were contemporary with Wertheimer, were unable to design teaching and training methods that promoted insight-based transfer that was superior to control conditions (Hilgard, Edgren, & Irvine, 1954; Katona, 1940). In a subsequent article Hatano and Inagaki (1986) made a similar proposal and distinguished between routine (mechanical) and adaptive (flexible and creative) expertise. Quite recently Morrison and Fletcher (2002) introduced the concept of *cognitive readiness*, which emphasizes the importance of similar adaptive abilities to display "competent performance in the complex and unpredictable environment of modern military operations" (pp. 1–3). Morrison and Fletcher (2002) emphasize the value of the ability to transfer trained performance to new situations, to generate flexible and creative problem solutions, to encounter challenges, and to implement action plans based on the generated solutions.

In this chapter I will review evidence from research on expertise and expert performance and its relevance to assertions made about the development and training of adaptive expertise and cognitive readiness. I will describe the expert-performance approach and how research within this theoretical framework has uncovered some

K.A. Ericsson (✉)
Department of Psychology, Florida State University, 1107 West Call Street,
P.O. Box 3064301, Tallahassee, FL 32306-1270, USA
e-mail: ERICSSON@PSY.FSU.EDU

H.F. O'Neil et al. (eds.), *Teaching and Measuring Cognitive Readiness*,
DOI 10.1007/978-1-4614-7579-8_10,

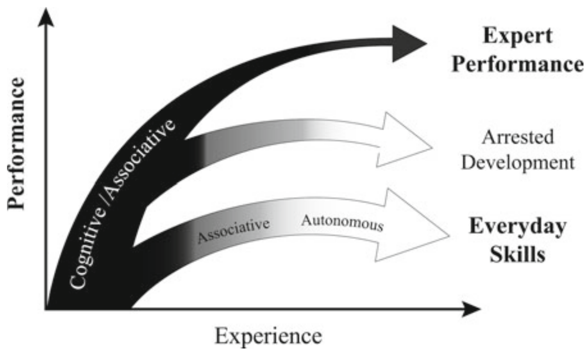


Fig. 10.1 An illustration of the qualitative difference between the course of improvement of expert performance and of everyday activities. The goal for everyday activities is to reach as rapidly as possible a satisfactory level that is stable and "autonomous." After individuals pass through the "cognitive" and "associative" phases they can generate their performance virtually automatically with a minimal amount of effort (see the *gray*/*white* plateau at the bottom of the graph). In contrast, expert performers counteract automaticity by developing increasingly complex mental representations to attain higher levels of control of their performance and will therefore remain within the "cognitive" and "associative" phases. Some experts will at some point in their career give up their commitment to seeking excellence and thus terminate regular engagement in deliberate practice to further improve performance which results in premature automation of their performance (adapted from "The scientific study of expert levels of performance: General implications for optimal learning and creativity" by K. A. Ericsson in *High Ability Studies*, *9*, p. 90. Copyright 1998 by European Council for High Ability)

generalizable insights into the structure and acquisition of expert performance. I will then discuss the proposals for adaptive expertise and cognitive readiness in the context of these findings. Let me first start by giving a brief historical background on research on skill and expertise that gives a backdrop to proposals for adaptive expertise and cognitive readiness.

10.1 Brief Historical Background for Modern Research on Expertise and Skill Acquisition

A review of research on skilled and expert performance reveals several distinctive aspects of expertise, skills, and superiority of performance. Ever since the pioneering research by Bryan and Harter (1897, 1899) on the development of telegraphic operator skill, the main attribute of attained skilled performance is its effortlessness. The most influential model of skill acquisition was proposed by Fitts and Posner (1967; Fitts, 1964). This model is consistent with the general experience of learning everyday activities such as driving a car, typing, or playing golf; people's primary goal is to reach an acceptable level of performance. During the first phase of learning of a new skill (Fitts & Posner, 1967), beginners try to understand the overall activity and try to complete various sequences of performance steps, as is illustrated in the first part of the lower arm of Fig. 10.1 (shown in black). As the individual gets more experience (the middle (gray) phase of learning), salient failures become rarer, and the individual is able to complete the steps of performance sequences more smoothly. The learners can generate their performance with less attention than was required in the beginning. Within a period of training and experience—frequently less than 50 h for most recreational activities—individuals attain an acceptable standard of performance, which can be elicited with dramatically reduced concentration and effort. At this advanced point, most learners no longer attempt to actively control their performance with the goal of making permanent modifications and improvements. During this last part of their skill learning, as is illustrated by the third (white) part of the lower arm in Fig. 10.1, performance typically reaches a stable plateau of performance that is merely maintained for months, years, and decades as long as the individual engages in regular activity.

This model influenced subsequent formalized models of skill acquisition. The theory of expertise proposed by Simon and Chase (1973) proposed that expertise is mediated by the acquisition of mental patterns (chunks) that permit the retrieval of appropriate actions previously stored in long-term memory. For instance, with more experience chess players will store a larger number of more complex patterns in memory and associate more sophisticated actions with these patterns. Hence, individual differences in chess playing will closely relate to the size and complexity of the accumulated library of pattern–action associations. Another related conceptualization of expertise was proposed by the Dreyfus brothers (Dreyfus & Dreyfus, 1986). They proposed a five-stage model of development of expertise that was closely linked to increased experience in the domain. The first stage involved learning the rules and concepts of the domain and then attempting to apply these rules in a deliberate sequential fashion. As a function of experience, the individual became increasingly proficient in producing adequate responses while going through the stages of the novice, the proficient, and finally reaching the expert stage, where actions were effortlessly generated based on an intuitive feeling. In later work Dreyfus and Dreyfus (1986) explicitly linked this progression to experience as they wrote that Some domains, such as driving a car, are simple and "almost all novices [beginners] can eventually reach the level we call expert" (p. 21), and more generally: "When things proceed normally, experts don't solve problems and don't make decisions, they do what normally works" (Dreyfus & Dreyfus, 1986, pp. 30–31).

In his influential paper on a theory of cognitive skill acquisition Anderson (1982) proposed a formal model within his ACT system consistent with Fitts and Posner's (1967) three stages of skill acquisition. The first declarative stage involves assembling the necessary knowledge (declarative knowledge) and procedures to be able to perform the tasks in a serial problem-solving mode. The second procedural stage permits integration of several consecutive actions into a single production and procedure, where knowledge is encoded and linked to the application conditions for the procedures—leading to speed-up of execution and reducing the need for step-by-step attentional control. In the third and final tuning stage, the productions are honed and adapted to the tasks. Anderson (1982, p. 403) argued that "this configuration of learning mechanisms described is involved in the full range of skill acquisition from language acquisition to problem solving to schema abstraction." As with most major theoretical advances, the proposal for theories of expertise (Simon & Chase, 1973) and cognitive skill acquisition (Anderson, 1982) was criticized, especially with respect to its ability to describe the full range of phenomena associated with expertise and skill acquisition.

10.2 Alternative Accounts of Acquisition of Skill and Expertise

There is considerable agreement that the Fitts and Posner (1967) model of skill acquisition and Simon and Chase (1973) theory of expertise explains some type of skill acquisition and expertise development. In the 1980s there were several investigators that criticized the assumption that with skill acquisition performance becomes increasingly specialized, automatic, and effortless. The most cited and perhaps most influential paper was "The two courses of expertise" written by Hatano and Inagaki (1986). They distinguished between two types of experts in the school setting, where one type was consistent with the Fitts-Posner description of attaining their expertise by solving a large number of problems, and where they "merely learn to perform a skill faster and more accurately, without constructing/enriching their conceptual knowledge" (p. 31). In contrast to these "routine" experts, Hatano and Inagaki (1986) proposed the existence of adaptive experts, who remained flexible and were guided by understanding even after extensive experience in solving problems. Consistent with the development of routine expertise Hatano and Inagaki (1986) found that extended practice with abacus calculation did not lead to increases in the school students' understanding of mathematics principles, such as carry operations. In order to develop understanding and the associated adaptive expertise, Hatano and Inagaki (1986) proposed that direct repetition should be avoided in favor of variability in problem situations. The teacher should strive to develop educational contexts in the school similar to those encountered in farming and cooking, where more experience appeared to lead to better understanding. It is problematic if teachers overemphasize accuracy in solving problems, because "when a procedural skill is performed primarily to obtain rewards, people are reluctant to take the risk of varying the skill, since they believe that the fast way is to rely on the 'conventional' version" (p. 34). Finally, in a culture that values understanding, students "are encouraged to try new versions of the procedural skills, even at the cost of efficiency to some extent; they are often requested to explain the appropriateness of the skill as well (mostly to others but sometimes to themselves)" (Hatano & Inagaki, 1986, p. 34). More recently, Bransford and Schwartz (1999) have made influential contributions to learning in the school context based on promoting the development of this type of adaptive expertise.

In the edited book "Toward a general theory of expertise: Prospects and limits" (Ericsson & Smith, 1991a), several contributors criticized the traditional models of the acquisition of skill and expertise. Ericsson and Smith (1991b) reviewed the history of research on expertise and found that de Groot's (1946/1978) original research on how expert chess players exhibited superior performance in winning games was replaced by a focus on a correlated and more easily studied superior performance, namely recalling briefly presented chess positions (Chase & Simon, 1973). Ericsson and Smith (1991b) proposed that we need to return to research on reproducibly superior performance and they proposed an alternative approach (based on de Groot's, 1946/1978 research on selecting the best move task) that captured representative performance in the laboratory to study its mediating mechanisms and how they were acquired. They reviewed the evidence and found that the acquired mechanisms mediating superior performance of experts differ qualitatively from those proposed by traditional theories of acquisition of skill and expertise. Among the invited commentators in the volume on expertise (Ericsson & Smith, 1991a), Holyoak (1991) and Salthouse (1991) made influential proposals and comments. Holyoak (1991) proposed that in order to account for the adaptive type of expertise proposed by Hatano and Inagaki (1986), one would have to develop a third generation of systems based on connectionism, namely symbolic connectionism. He argued that this type of system would be able to explain a number of observations on expertise, such as some flexibility in experts' performance, the decoupling of memory performance from expertise, the lack of uniform improvement with continued practice, and many more. In his commentary, Salthouse (1991) discusses the kind of limitations, such as information processing capacities, that an expert has to overcome to attain a superior level of performance in a particular domain. Salthouse (1991) recommended careful analysis of the mechanisms acquired by experts to circumvent general limits on their performance and argues that their structure will provide insights into the generalizability of the acquired superior performance.

The research on expertise raised doubts about the sufficiency of the modal theories of expertise that proposed automation in response to extensive experience with the task. This research demonstrating considerable cognitive mediation and its associated control, thus supporting Hatano and Inagaki's (1986) hypothesis of differentiating "routine" from adaptive expertise. At the same time, this research cautioned the optimism of Hatano and Inagaki (1986) in developing adaptive expertise.

It would be wonderful if one could develop adaptive methods for developing expertise in students. Then based on their deep understanding of the domains of knowledge, they would be able to produce creative solutions to unfamiliar problems and challenges. Research on the performance of experts shows that under some circumstances it is possible to attain reproducibly superior performance after extended engagement in appropriate forms of practice. However, the conditions of practice that produce superior performance differ markedly from those proposed by Hatano and Inagaki (1986), as I will show later in this chapter.

10.3 Transfer and the Recent Interest in Cognitive Readiness in Professional Expertise

Most of the research on skill acquisition and learning in school environments shows that the acquired performance is quite specific to the tasks used in actual training activities (see Detterman, 1993; Singley & Anderson, 1989, for general reviews). Consistent with the criticisms of "routine" expertise, students seem to master just the types of problems that they have trained on. For example, vanLehn and van der Sande (2009) found that physics students were able to solve standard textbook problems by solving equations but were surprisingly poor at reasoning about similar problems at a qualitative level. More importantly, there appears to be a gap between general education in schools on basic mathematical and verbal skills, as well as sciences, history and foreign languages, and the necessary preparation for work and professional activity. Even professional schools for engineering, medicine, business, and law have acknowledged the marked transition from theoretical learning in the school environment to real job performance and establishment of a professional career. Evidence for the lack of connection between professional schools and actual professional success is provided by the finding that grades in professional schools are almost unrelated to professional success as measured by salary—Roth and Clarke's (1998) meta-analysis found a correlation of 0.15 with starting salary and only a correlation of 0.05 with salary growth. These findings raise questions about the nature of the factors that lead to professional success. Traditional criteria for identifying experts were social reputation, completed education, accumulated accessible knowledge, and length of experience in a domain (over 10 years) (Chi, 2006; Chi, Glaser, & Farr, 1988), but in studies where objective performance has been measured there is frequently not a significant correlation between amount of professional experience or professional training and performance (Ericsson, 2004, 2009; Ericsson & Lehmann, 1996; Ericsson, Whyte, & Ward, 2007). Sometimes there is even a significant negative correlation between length of experience and objective performance (Choudhry, Fletcher, & Soumerai, 2005). For example, research has shown that highly experienced computer programmers' performance on programming tasks is not always superior to that of computer science students, and physics professors from a first-rate research university were not always consistently superior to students on introductory physics problems (see Ericsson & Lehmann, 1996, for a review). In a recent review of political judgment, Tetlock (2005) compared predictions from hundreds of experts in different fields with well-informed nonexperts and were able to dispel the myth that experts' forecasts are superior.

In light of the evidence that extended experience appears only to strengthen behavior in typical situations, as predicted by the Fitts and Posner (1967) model of skill acquisition, it seems natural that planners would be interested in developing more general skills as well as a cognitive readiness to respond "in complex and unpredictable environments" (Morrison & Fletcher, 2002). Morrison and Fletcher (2002) reviewed the literature to find studied characteristics that would be likely to lead to more generalizable, flexible, and creative performance. They then go on to discuss methods for training these characteristics, such as transfer of training, metacognition, problem solving, decision making, and mental flexibility and creativity. Then, by analyzing the available evidence from the primarily laboratory studies that Morrison and Fletcher (2002) cite, it may be possible to apply these ideas and training principles and gain meaningful improvement in cognitive readiness.

In this chapter I will discuss what we know about how expert and very high levels of measurable performance is attained and then comment on the proposals generated from the adaptive expertise perspective and their relevance for cognitive readiness.

10.4 The Expert-Performance Approach

Whereas the traditional experimental psychologists seek to identify the phenomena in their most basic and pure form and then generalize findings to more complex phenomena, the expert-performance approach proceeds in the reverse direction. Researchers pursuing the expert-performance approach seek to identify experts and other individuals who can consistently reproduce the superior performance in question. If we are interested in finding superior performance in dealing with unexpected phenomena, we might want to look for outstanding performers in competitive domains, such as chess and sports. The expert-performance approach focuses on reproducibly superior performance on representative, authentic tasks in their field (Ericsson & Smith, 1991b). For example, the focus might be on the processes of diagnosis and treatment of patients leading to superior outcomes for patients, on the consistent selection of the best moves for chess positions that lead to winning chess games, or on superior performance in music and sport competitions that lead to medals and prizes. The first step in expert-performance approach requires that researchers be able to capture the reproducibly superior performance on some representative tasks of the domain of expertise, and then be able to examine this superior performance with laboratory methods, as described in the following section.

10.4.1 Capturing Reproducibly Superior Performance Under Standardized Conditions

In everyday life, experts encounter different challenges under different conditions, which makes it virtually impossible to measure levels of performance of different experts. For example, one doctor may treat two patients with problems that require risky and complex treatment, such as major surgery, whereas another may treat six patients with problems that can be successfully be treated with standard treatments, such as medications. One manager has to resolve several serious interpersonal conflicts during a restructuring of a firm and another manager merely has to guide an enthusiastic team. Unless individuals are presented with the same or comparable situations, it will be very difficult to measure individual differences in performance.

In a pioneering study of world-class chess players, de Groot (1946/1978) designed representative tasks that captured their superior performance. The ultimate measure of success in chess is becoming the winner at chess tournaments, where players compete against each other in matches that last for several hours. There are statistical methods for measuring a chess player's skill on an interval scale of chess ability from outcomes of 20–50 tournament matches (roughly corresponding to 40–100 h of play). Using a pioneering approach, de Groot (1946/1978) developed a method to elicit the cognitive processes distinguishing better chess player from weaker ones that could be assessed in less than half an hour. He identified particular chess positions (selected from past games between chess masters), where the selection of the best next move was critical to the outcome of the chess match. He then presented these positions to chess players during an individual testing session and asked them to "think aloud" while they generated the best possible next move for each of the positions. De Groot (1946/1978) demonstrated that world-class players reliably found the best moves, whereas skilled club players did not. In this study the best move for a position was determined by months of continued analysis of that particular chess position by chess masters. Subsequent research with large groups of chess players differing widely in skill has shown that the selection of the best move for selected chess positions is highly correlated with the tournament ratings (Ericsson, Patel, & Kintsch, 2000; van der Maas & Wagenmakers, 2005). When performance on 20 selection tasks is aggregated, the resulting score is highly correlated with chess ratings—thus, it is possible to measure a chess player's skill after less than 15 min of testing. Of particular interest, de Groot (1946/1978) was able to identify how the thought processes of world-class players displayed qualitative differences from those of skilled club players by analyzing their think aloud protocols from the selection tasks.

Based on de Groot's paradigm, it is possible to apply it to the study of experts' performance in a given domain of expertise, such as music, sports, or medicine. The first step involves the identification of naturally occurring situations in a given domain of expertise that require immediate action and where successful performance captures the essence of expertise in the associated domain. For example, in their everyday routines doctors will encounter situations where they have to assess the symptoms of patients for immediate diagnosis and treatment. Once such situations from the studied domain of expertise have been selected and appropriate courses of action identified, it is then possible to reproduce them with appropriate context and an immediate demand for action under standardized conditions for all tested individuals.

The representative tasks are designed to simulate task demands and situations that a performer might encounter. For example, if chess masters consistently select the best move for presented positions in the laboratory, then they should be able to select the best moves during chess games and thus beat most opponents. The representative tasks in Fig. 10.2 have been found to capture the essence of the associated expertise in the three respective domains of expertise. To measure chess skill, players at different skill levels are asked to generate the best move for identical positions where a single superior move is available (as determined by new chess-playing programs that are superior to the best human players for tactical problems). To measure typing speed, typists are asked to copy as much of the presented material as possible during a fixed time period. When musicians are instructed to play an unfamiliar piece of music and then play it again as similar to the original performance as is possible, expert musicians are able to reproduce a given performance with less variability than less skilled musicians (Krampe & Ericsson, 1996).

Across domains of expertise with objective measurements, a couple of relations between expert performance and experience emerge (Ericsson, 2006b, 2009). First, markedly superior performance is only attained after extensive experience in a given domain. Second, only certain types of experience lead to improved performance. In fact, many thousands of hours of particular types of practice and training are necessary for reaching the highest levels of performance.

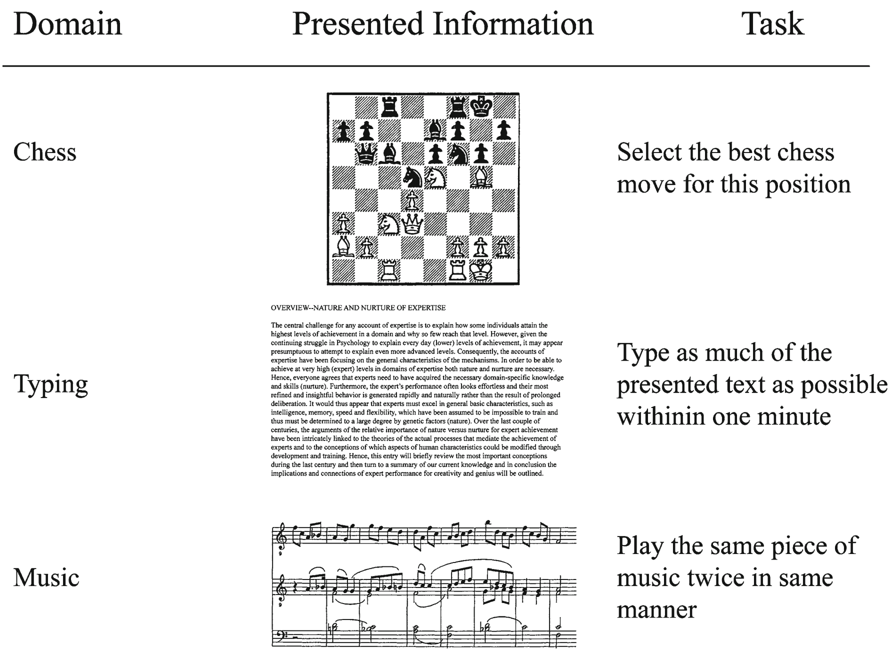


Fig. 10.2 Three examples of laboratory tasks that capture the consistently superior performance of domain experts in chess, typing and music (from "Expertise," by K. A. Ericsson and Andreas C. Lehmann, 1999, *Encyclopedia of Creativity*. Copyright by Academic Press)

10.4.2 The Necessity of Domain-Specific Experience for Attaining Reproducibly Superior Performance

Expert performance is only attained after extended engagement in domain-related activities in the corresponding domain (Ericsson, 2006b, 2009; Ericsson & Lehmann, 1996). The level of performance of experts can be sketched as a function of years of experience in the domain, as follows. First, longitudinal measurements of performance of experts reveal that all individuals improve gradually, as illustrated in Fig. 10.3. When performance is measured using the same adult standards, there appears to be no objective evidence that a child or adult is able to exhibit a high level of performance without any relevant prior experience and training (Ericsson, Nandagopal, & Roring, 2009; Ericsson, Roring, & Nandagopal, 2007a, 2007b). When the performances of child prodigies in music and chess are measured using adult standards, these children show gradual, steady improvement over time. Second, elite performance keeps improving beyond the age of physical maturation—the late teens in industrialized countries. Experts' peak performances during their careers are nearly always attained in adulthood—many years, and even decades, after initial introduction to activities in the domain, as illustrated in Fig. 10.3. The age at which performers typically reach their highest level

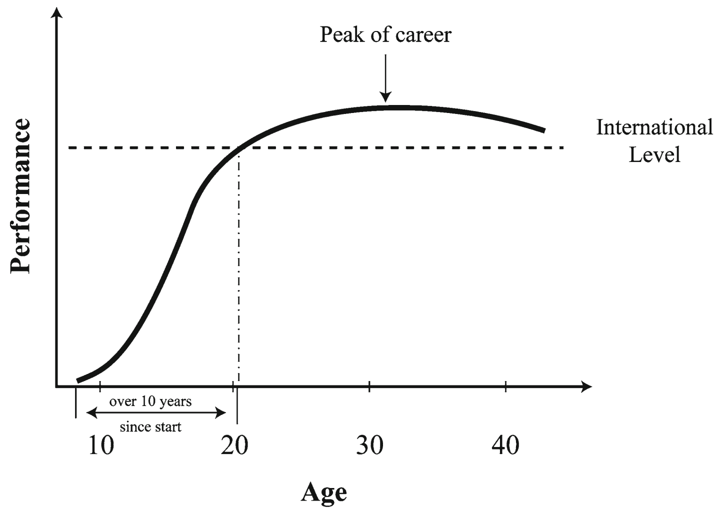


Fig. 10.3 An illustration of the gradual increases in expert performance as a function of age, in domains such as chess. The international level, which is attained after more than around 10 years of involvement in the domain, is indicated by the *horizontal dashed line* (from "Expertise," by K. A. Ericsson and Andreas C. Lehmann, 1999, *Encyclopedia of Creativity*. Copyright by Academic Press)

of performance in many vigorous sports is the mid- to late 1920s. For the arts and science, it is a decade later, in the 1930s and 1940s (see Schulz & Curnow, 1988; Simonton, 1997, 2004, for reviews). The continued increases of experts' performances continue past physical maturity, which shows that the highest level of performance must be the result of further learning and physiological adaptation due to training and experience.

Finally, vast experience in domain-related activities is necessary to attain expert performance. Investigators have shown that everyone, even the most talented individuals, need to invest 10 years of active engagement in a domain (10-year rule) to reach an expert level of performance (Bryan & Harter, 1899) and to win in international competitions (Simon & Chase, 1973). Subsequent reviews have found that the 10-year rule extends to international level success in music composition (Hayes, 1981), as well as to sports, science, and the arts (Ericsson, Krampe, & Tesch-Römer, 1993). Finally, outstanding scientists and authors normally published their first work at around age 25, and their best work follows around 10 years later (Raskin, 1936; Skirbekk, 2008).

10.5 From Experience to Designed Practice

Most active participants in domain-related activities seem satisfied with reaching a merely acceptable level of performance, such as amateur tennis players and golfers. Once an acceptable level has been reached, a stable performance can often be maintained with minimal effort for years and even decades. In contrast, high-level performers in the arts, science, and sports have a fundamentally different developmental history than their less accomplished peers (Bloom, 1985a, 1985b). The parents of Bloom's future elite performers helped their children from an early age to get help from teachers and get access to the best training resources.

At the same time, simply having access to the best training environments is not sufficient to produce the very best performers, because there are individual differences even among the select individuals practicing in these environments. In a study of expert violinists, Ericsson et al. (1993) found that the best violinists spent more time per week on activities that had been specifically designed to improve performance, which we named "deliberate practice." Expert musicians working in a solitary setting in order to master-specific goals determined by their music teacher is a good example of goal-directed efforts to reach a higher level of performance. At a particular lesson, the musicians are instructed by their teachers to improve some aspect of their playing. A week later, and after considerable amount of solitary practice on relevant exercises, the musicians are now able to play at the requested level. When the same expert violinists were interviewed to derive an estimate of the amount of solitary (deliberate) practice during their musical development, the most accomplished musicians were found to have spent significantly more time engaging in deliberate practice during their development.

The central assumption of the theoretical framework of the expert-performance approach (Ericsson, 1996, 2004; Ericsson et al., 1993) is that expert performance can be described as a sequence of states associated with increasingly higher levels of performance. In the acquisition of music performance, pianists acquire increasingly complex playing techniques in consecutive fashion with the most difficult techniques last. Similar findings have been made in sports, where the performance of individual athletes improves over their career (Ericsson, 2007a; Ericsson & Towne, 2010).

Tasks suitable for deliberate practice will depend on the particular training goal and the individuals' preexisting skills for monitoring and controlling their performance (Ericsson, 2006a, 2007a, 2007b). For example, a chess player that is presented with the task of finding the best move for a specific chess position will use their acquired representations to find the best move by planning and reasoning. If they find out that they failed to generate the best move, then they will engage in analysis to figure out why they did not find the best move to be able to avoid similar mistakes in the future.

The main assumption is that an individual's performance on a training task will vary as a function of focus of attention, type of strategy, and many other situational factors. If one wants to produce one's highest current performance consistently and, in particular, go beyond one's current maximal performance, one has to identify the optimal factors. Consequently, any type of deliberate practice is designed to maximize performance gains by allowing the performer to be fully rested at the start of the deliberate practice activity. Furthermore, the performers should also be fully prepared for initiation of the task, be given immediate feedback from the outcome and then allowed to repeat the same or similar task with gradual modifications. Performing the task under these optimal conditions is more effective than performing the associated task only when it occurs, often unexpectedly, within the natural context of performance. Hence, once a performance improvement has been attained, part of the ensuing practice gradually embeds the trained task in its natural context with regular time constraints and less predictable occurrences. For example, consider an amateur tennis player who misses a volley at the net. The tennis game will continue until some time later a similar situation with the need to return the ball with a volley happens unexpectedly—with a similar problem for the player. Compare this type of "on-the-job learning" with a session with a tennis coach. The tennis coach would arrange training situations where the player would stand at the net and be able to anticipate the volley. With mastery of the easy volleys, the coach can increase the difficulty of the shots and eventually embed volley shots into the rallies. It is easy to see that a few hours of this type of training would improve the player's volley more than tens or hundreds of hours of regular tennis play against other amateurs.

The focus on and the goal of improving performance sets deliberate practice apart from both mindless, routine performance and playful games as the latter types of activities would, if anything, merely strengthen the currently inadequate cognitive mediating mechanisms rather than modify them to allow increased performance. In direct contrast, aspiring expert performers never allow their performance to be fully automated but continue to seek out, with the help of their teachers, new training activities that require them to engage in problem solving to alter their cognitive representations that allow them to keep improving the mechanisms mediating performance, as illustrated by the upper arm in Fig. 10.1. Some performers, who initially aspired to reach expert levels of performance, will eventually stop pushing their limits after months or even years of training (Ericsson, Perez, et al., 2009). The performance level at the end of training will be eventually automated and their development of performance will be prematurely arrested at that intermediate level as illustrated by the middle arm in Fig. 10.1.

In chess Charness and his colleagues (Charness, Krampe, & Mayr, 1996; Charness, Tuffiash, Krampe, Reingold, & Vasyukova, 2005) have found that the amount of solitary chess study was the best predictor of performance at chess tournaments, and when this type of deliberate practice was statistically controlled, there was no reliable benefit from playing chess games. During solitary practice, aspiring chess experts study published games and try to play against the masters by trying to select the best move for each step of the game. If they successfully predict the next move by chess masters, then they match the level of play—at least for that particular move. If they fail to predict the move selected by the chess master, they study the position more closely to discover the reasons for the master's move. By this type of immediate feedback, players are able to make corrections to their representations and their search for alternative combinations of chess moves. By figuring out why they made a mistake or selected an inferior move, they are able to make changes that permit them to avoid that particular mistake as well as mistakes for other related chess positions. Similar findings have been made in other domains. For example, in a study of district, national, and professional dart players Duffy, Baluch, and Ericsson (2004) found that solitary deliberate practice was closely related to performance, whereas the amount of social dart activities did not predict performance. For a review of various types of deliberate practice, see Ericsson (2006b).

10.6 When Can Professionals Learn and Improve Their Performance

Most professionals, such as doctors, nurses, stock brokers, and accountants, do not receive immediate feedback on their objective performance and this fact may explain why many doctors do not spontaneously adopt the best treatment methods for their patients (Davis, 2009) and spend a rather modest amount of time engaged in deliberate practice and effortful training to improve and maintain their skills and keep updated on the most recent effective treatments.

The greatest obstacle for engaging in deliberate practice while working is the lack of immediate objective feedback. Frequently, the accuracy of decisions about medical treatments or investments in stocks or securities may not be apparent for weeks or months until consequences of the decisions can be observed. In work environments, typically a more experienced individual, such as a supervisor, will give immediate feedback. However, this feedback is unfortunately only as good as the objective performance of the supervisor. For example, the training of new radiologists lets medical interns make their diagnoses of incoming mammograms and then they receive feedback on their diagnoses by the senior radiologists reviewing the same X-rays. However, the accuracy of the senior radiologists is only around 70 % for correctly diagnosing malignant growths, so the radiology interns can only slowly reach a level matching their best teachers.

Another challenge to learning in professional domains is that difficult and challenging situations occur quite infrequently. Frequently encountered tasks and problems are often handled reasonably well by most professionals, and the main individual differences arise when dealing with infrequent and urgent problems, such as emergency situations for airline personnel and emergency room staff (Ericsson, Perez, et al., 2009). To train personnel to respond appropriately in these situations, it is necessary to conduct training in simulators where emergency situations can be created and thus experienced. In an important study, McKinney and Davis (2004) found that when the expert pilots had practiced the same emergency situation in the simulator prior to the emergency event, they were reliably more successful dealing with an actual event and thus the specific training reduced the probability of negative outcomes for pilot and plane.

With an increased interest in the life-long development of professional performance, there should be a greater interest in designing training environments with challenging relevant situations that require immediate action and that can provide feedback and opportunities for repeated encounters of the same and similar tasks.

10.7 Summary and Remarks on Adaptive Expertise and Cognitive Readiness

The research on the acquisition of expert performance reveals some necessary conditions for continued learning. Once individuals have reached an acceptable level of performance in a domain, where they receive insufficient feedback on errors, such as errant strokes in golf and tennis, improvements in performance do not occur automatically or without discernable causes. Most types of habitual professional and everyday activities such as driving a car, typing, or carrying out familiar task, tend to automate the associated behavior to minimize the effort required for execution of the desired performance. Continued improvements can be linked to individuals actively seeking feedback by teachers or mentors and to design training environments where they can gradually refine and generalize changes into the control mechanisms of their behavior and performance. The essence of deliberate practice is that individuals seek out challenges that go beyond their current level of reliable performance—ideally in a safe learning context that allows immediate feedback and gradual refinement by repetition. These learning environments can be viewed as scaffolds that facilitate attainment of a higher level of performance. Later the scaffolds can be gradually eliminated so performance can be embedded and elicited in the natural environments in the domain of expertise.

These learning environments seem both to resemble and to differ from those advocated by Hatano and Inagaki (1986) for developing adaptive expertise. The most important factor of supportive environments, according to Hatano and Inagaki (1986), is the variability in encountered tasks and situations and the "built-in randomness" (p. 34). The expert-performance approach agrees with Hatano and Inagaki (1986) that "repeated application of a procedure is unlikely to lead to adaptive expertise" (p. 33). They criticize the default practice of "blind" repetition in the schools and propose that one should mimic the natural environments, such as growing plants or cooking food, to develop adaptive expertise. However, the review earlier in this chapter showed that mere extended engagement in everyday activities, such as driving, most likely gardening, and cooking, does not lead to generalizable improvements in performance. Increased experience appears to lead to assimilation of new experiences into established automated routines. The primary direction of the expert-performance approach is toward continued experience in natural environments and the need for goal-directed training to improve specific aspects of performance in special training environments. Hatano and Inagaki's (1986) second factor promoting adaptive expertise concerns the negative role of immediate external rewards. These rewards are assumed to push students toward using safe strategies with reliable short-term outcomes at the expense of more creative methods leading to long-term improvements. This point is similar to the distinction between work and deliberate practice made by Ericsson et al. (1993). In work situations, individuals will seek reliable strategies with proven effectiveness whereas the improvements in performance will occur only when individuals can seek out challenging situations, where failure is likely, in order to successively master these conditions. The final factor promoting adaptive expertise according to Hatano and Inagaki (1986) concerns the emphasis on understanding by the learning community. The emphasis on understanding might be appropriate as a reaction to the mindless memorization of equations and solution principles in many school systems. The expert-performance approach, however, focuses more on the development of representations for planning, reasoning, and evaluation of one's actions, so performers can self-assess their weaknesses and design appropriate training activities to keep improving their performance even in the absence of teachers and coaches. The differences between the adaptive expertise emphasis (Hatano & Inagaki, 1986) and the expert-performance approach becomes less clear when one considers the context of their primary application. Hatano and Inagaki (1986) discussed how the learning can be changed in the schools, whereas the expert-performance approach seeks to understand how individuals attain reproducibly superior performance for representative tasks that capture the essence of expertise in a domain or professional activity.

The findings from the expert-performance approach have implications for proposals to develop cognitive readiness (Morrison & Fletcher, 2002). Most critically, the expert-performance approach cannot be applied to a performance phenomenon unless it is possible to measure the associated objective performance. It will be difficult to measure individuals' performance in a wide range of situations with unexpected outcomes. Perhaps the first step would be to specify how to measure the ability to respond appropriately in a promising type of situation with *low* predictability. Unless it is possible to find individuals with a high level of performance for this test, or finding individuals, who can attain a high level of performance after training, it seems pointless to attempt to search for individuals who can excel in many different types of situations. The findings of the expert-performance approach suggest that high levels of performance in a given domain is mediated by gradually acquired complex integrated systems of representations for the execution, monitoring, planning and analysis of performance. As a consequence acquisition of skilled performance requires an orderly and deliberate approach. Deliberate practice requires the designing of training tasks that capture the conditions of representative situations, where the performer can refine and change the complex cognitive mechanisms that mediate (generate) superior performance. Improvements always involve the performers' preexisting mechanisms. The tight interrelation between representations that monitor and generate performance minimizes the risk of unwanted side-effects from modifications. However, the complex integration of the mechanisms mediating superior expert performance makes it impossible to isolate distinct processes of problem solving, decision making, and reasoning. In fact, the principal challenge of professional skill acquisition appears to be in developing representations that coordinate each of the following: selection of actions, monitoring, control of ongoing performance, and incremental improvements to performance.

Consistent with the proposal for cognitive readiness (Morrison & Fletcher, 2002), the findings from the expert-performance approach show that superior performance is associated with superior representation of the current situation (situational awareness, memory), with superior ability to reason and evaluate (metacognition, decision making, problem solving, and mental flexibility), and superior performance in previously unencountered situations (transfer of learning). At the same time the expert-performance approach emphasizes the need to build representations that permit integrated generation of appropriate behavior in a class of possible situations—the development and refinement of these representations takes place during deliberate practice in representative situations when generalizations are crafted in response to processing of feedback. One conceptualization that would be consistent with the development of a "general" ability to respond to a wide range of situations would be to classify expert chess performance as the ability to generate superior moves in a broad range of different types of chess situations—recall that the number of different chess positions are estimated to exceed 10^{40} (Shannon, 1950). More research studying how expert chess players develop their ability to respond to such a large number of different positions might inform how similar skills could be developed for military personnel to react in an appropriate manner in situations generated by an adversary.

The research on expert levels of achievement in traditional and professional domains of expertise will need to measure and study superior performance in specific domains. There are, however, findings that are likely to generalize across domains, such as the characteristics of ideal training environments, and methods for fostering motivation by providing both emotional support and attainable training tasks of a suitable difficulty level. Furthermore, it is likely that research on the effective development of skills in one domain, such as bowing technique in violin playing, can inform newer types of activities, such as robotic surgery. The integrated knowledge about how different types of skills can be effectively acquired has implications not just for elite performance but also for training students in general education, as some of them might be interested in striving to reach expert levels in professional domains of expertise. By examining how the prospective expert performers attain the beginning levels of achievement, we should be able to develop practice environments that help students to attain the fundamental representations and the self-regulatory skills that would give them the ability to continue to increasingly higher levels of achievement.

Acknowledgements This research was supported by the FSCW/Conradi Endowment Fund of Florida State University Foundation. The author wants to thank Len Hill and Jerad Moxley, for the helpful comments on earlier drafts of the chapter. The work reported herein was partially supported by a grant from the Office of Naval Research, Award Number N000140810126. The findings and opinions expressed here do not necessarily reflect the positions or policies of the Office of Naval Research.

References

Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review, 89*, 369–406.

Bloom, B. S. (1985a). Generalizations about talent development. In B. S. Bloom (Ed.), *Developing talent in young people* (pp. 507–549). New York: Ballantine Books.

Bloom, B. S. (Ed.). (1985b). *Developing talent in young people*. New York: Ballantine Books.

Bransford, J. D., & Schwartz, D. L. (1999). Rethinking transfer: A simple proposal with multiple implications. *Review of Research in Education, 24*, 61–100.

Bryan, W. L., & Harter, N. (1897). Studies in the physiology and psychology of the telegraphic language. *Psychological Review, 4*, 27–53.

Bryan, W. L., & Harter, N. (1899). Studies on the telegraphic language: The acquisition of a hierarchy of habits. *Psychological Review, 6*, 345–375.

Charness, N., Krampe, R., & Mayr, U. (1996). The role of practice and coaching in entrepreneurial skill domains: An international comparison of life-span chess skill acquisition. In K. A. Ericsson (Ed.), *The road to excellence: The acquisition of expert performance in the arts and sciences, sports, and games* (pp. 51–80). Hillsdale, NJ: Erlbaum.

Charness, N., Tuffiash, M. I., Krampe, R., Reingold, E., & Vasyukova, E. (2005). The role of deliberate practice in chess expertise. *Applied Cognitive Psychology, 19*, 151–165.

Chase, W. G., & Simon, H. A. (1973). The mind's eye in chess. In W. G. Chase (Ed.), *Visual information processing* (pp. 215–281). New York: Academic.

Chi, M. T. H. (2006). Two approaches to experts' characteristics. In K. A. Ericsson, N. Charness, P. Feltovich, & R. R. Hoffman (Eds.), *Cambridge handbook of expertise and expert performance* (pp. 21–38). Cambridge, UK: Cambridge University Press.

Chi, M. T. H., Glaser, R., & Farr, M. J. (Eds.). (1988). *The nature of expertise*. Hillsdale, NJ: Erlbaum.

Choudhry, N. K., Fletcher, R. H., & Soumerai, S. B. (2005). Systematic review: The relationship between clinical experience and quality of health care. *Annals of Internal Medicine, 142*, 260–273.

Davis, D. A. (2009). How to help professionals maintain and improve their knowledge and skills: Triangulating best practices in medicine. In K. A. Ericsson (Ed.), *The development of professional expertise: Toward measurement of expert performance and design of optimal learning environments* (pp. 180–202). Cambridge, UK: Cambridge University Press.

de Groot, A. (1978). *Thought and choice and chess.* The Hague: Mouton. (Original work published 1946)

Detterman, D. K. (1993). The case for prosecution: Transfer as an epiphenomenon. In D. K. Detterman & R. J. Sternberg (Eds.), *Transfer on trial: Intelligence, cognition, and instruction* (pp. 39–67). Stamford, CT: Ablex.

Dreyfus, H. L., & Dreyfus, S. E. (1986). *Mind over machine: Power of human intuitive expertise in the era of the computer*. New York: Free Press.

Duffy, L. J., Baluch, B., & Ericsson, K. A. (2004). Dart performance as a function of facets of practice amongst professional and amateur men and women players. *International Journal of Sport Psychology, 35*, 232–245.

Ericsson, K. A. (1996). The acquisition of expert performance: An introduction to some of the issues. In K. A. Ericsson (Ed.), *The road to excellence: The acquisition of expert performance in the arts and sciences, sports, and games* (pp. 1–50). Hillsdale, NJ: Erlbaum.

Ericsson, K. A. (2004). Deliberate practice and the acquisition and maintenance of expert performance in medicine and related domains. *Academic Medicine, 10*, S1–S12.

Ericsson, K. A. (2006a). The influence of experience and deliberate practice on the development of superior expert performance. In K. A. Ericsson, N. Charness, P. Feltovich, & R. R. Hoffman (Eds.), *Cambridge handbook of expertise and expert performance* (pp. 685–706). Cambridge, UK: Cambridge University Press.

Ericsson, K. A. (2006b). Protocol analysis and expert thought: Concurrent verbalizations of thinking during experts' performance on representative task. In K. A. Ericsson, N. Charness, P. Feltovich, & R. R. Hoffman (Eds.), *Cambridge handbook of expertise and expert performance* (pp. 223–242). Cambridge, UK: Cambridge University Press.

Ericsson, K. A. (2007a). Deliberate practice and the modifiability of body and mind: Toward a science of the structure and acquisition of expert and elite performance. *International Journal of Sport Psychology, 38*, 4–34.

Ericsson, K. A. (2007b). Deliberate practice and the modifiability of body and mind: A reply to the commentaries. *International Journal of Sport Psychology, 38*, 109–123.

Ericsson, K. A. (2009). Enhancing the development of professional performance: Implications from the study of deliberate practice. In K. A. Ericsson (Ed.), *The development of professional expertise: Toward measurement of expert performance and design of optimal learning environments* (pp. 405–431). New York: Cambridge University Press.

Ericsson, K. A., Krampe, R. T., & Tesch-Römer, C. (1993). The role of deliberate practice in the acquisition of expert performance. *Psychological Review, 100*, 363–406.

Ericsson, K. A., & Lehmann, A. C. (1996). Expert and exceptional performance: Evidence on maximal adaptations on task constraints. *Annual Review of Psychology, 47*, 273–305.

Ericsson, K. A., Nandagopal, K., & Roring, R. W. (2009). An expert-performance approach to the study of giftedness. In L. Shavinina (Ed.), *International handbook of giftedness* (pp. 129–153). Berlin: Springer.

Ericsson, K. A., Patel, V. L., & Kintsch, W. (2000). How experts' adaptations to representative task demands account for the expertise effect in memory recall: Comment on Vicente and Wang (1998). *Psychological Review, 107*, 578–592.

Ericsson, K. A., Perez, R., Eccles, D. W., Lang, L., Baker, E., Bransford, J., et al. (2009). The measurement and development of professional performance: An introduction to the topic and the organization of this book. In K. A. Ericsson (Ed.), *The development of professional expertise: Toward measurement of expert performance and design of optimal learning environments* (pp. 1–24). New York: Cambridge University Press.

Ericsson, K. A., Roring, R. W., & Nandagopal, K. (2007a). Giftedness and evidence for reproducibly superior performance: An account based on the expert-performance framework. *High Ability Studies, 18*, 3–56.

Ericsson, K. A., Roring, R. W., & Nandagopal, K. (2007b). Misunderstandings, agreements, and disagreements: Toward a cumulative science of reproducibly superior aspects of giftedness. *High Ability Studies, 18*, 97–115.

Ericsson, K. A., & Smith, J. (Eds.). (1991a). *Toward a general theory of expertise: Prospects and limits*. Cambridge, UK: Cambridge University Press.

Ericsson, K. A., & Smith, J. (1991b). Prospects and limits in the empirical study of expertise: An introduction. In K. A. Ericsson & J. Smith (Eds.), *Toward a general theory of expertise: Prospects and limits* (pp. 1–38). Cambridge, UK: Cambridge University Press.

Ericsson, K. A., & Towne, T. J. (2010). Expertise. *Wiley Interdisciplinary Reviews: Cognitive Science, 1*, 404–416.

Ericsson, K. A., Whyte, J., & Ward, P. (2007). Expert performance in nursing: Reviewing research on expertise in nursing within the framework of the expert-performance approach. *Advances in Nursing Science, 30*, E58–E71.

Fitts, P. (1964). Perceptual-motor skill learning. In A. W. Melton (Ed.), *Categories of human learning* (pp. 243–285). New York: Academic.

Fitts, P., & Posner, M. I. (1967). *Human performance*. Belmont, CA: Brooks/Cole.

Hatano, G., & Inagaki, K. (1986). Two courses of expertise. In H. Stevenson, H. Azuma, & K. Hakuta (Eds.), *Child development and education in Japan* (pp. 262–272). New York: W.H. Freeman.

Hayes, J. R. (1981). *The complete problem solver*. Philadelphia, PA: Franklin Institute Press.

Hilgard, E. R., Edgren, R. D., & Irvine, R. P. (1954). Error in transfer following learning with understanding: Further studies with Katona's card-trick experiments. *Journal of Experimental Psychology, 47*, 457–464.

Holyoak, K. J. (1991). Symbolic connectionism: Toward third-generation theories of expertise. In K. A. Ericsson & J. Smith (Eds.), *Toward a general theory of expertise: Prospects and limits* (pp. 301–335). Cambridge, UK: Cambridge University Press.

Katona, G. (1940). *Organizing and memorizing*. New York: Columbia University Press.

Krampe, R. T., & Ericsson, K. A. (1996). Maintaining excellence: Deliberate practice and elite performance in young and older pianists. *Journal of Experimental Psychology: General, 125*, 331–359.

McKinney, E. H., & Davis, K. J. (2004). Effects of deliberate practice on crisis decision performance. *Human Factors, 45*, 436–444.

Morrison, J. E., & Fletcher, J. D. (2002). *Cognitive readiness* (IDA Paper P-3735). Alexandria, VA: Institute for Defense Analyses.

Raskin, E. (1936). Comparison of scientific and literary ability: A biographical study of eminent scientists and letters of the nineteenth century. *Journal of Abnormal and Social Psychology, 31*, 20–35.

Roth, P. L., & Clarke, R. L. (1998). Meta-analyzing the relation between grades and salary. *Journal of Vocational Behavior, 53*, 386–400.

Salthouse, T. A. (1991). Expertise as the circumvention of human processing limitations. In K. A. Ericsson & J. Smith (Eds.), *Toward a general theory of expertise: Prospects and limits* (pp. 286–300). Cambridge, UK: Cambridge University Press.

Schulz, R., & Curnow, C. (1988). Peak performance and age among superathletes: Track and field, swimming, baseball, tennis, and golf. *Journal of Gerontology: Psychological Sciences, 43*, 113–120.

Shannon, C. E. (1950). Programming a computer for playing chess. *Philosophical Magazine, 41*, 256–275.

Simon, H. A., & Chase, W. G. (1973). Skill in chess. *American Scientist, 61*, 394–403.

Simonton, D. K. (1997). Creative productivity: A predictive and explanatory model of career trajectories and landmarks. *Psychological Review, 104*, 66–89.

Simonton, D. K. (2004). *Creativity in science: Chance, logic, genius, and zeitgeist*. Cambridge, UK: Cambridge University Press.

Singley, M. K., & Anderson, J. R. (1989). *The transfer of cognitive skill*. Cambridge, MA: Harvard University Press.

Skirbekk, V. (2008). Age and productivity potential: A new approach based on ability levels and industry-wide task demand. *Population and Development Review, 34*(Suppl), 191–207.

Tetlock, P. E. (2005). *Expert political judgment*. Princeton, NJ: Princeton University Press.

van der Maas, H. L. J., & Wagenmakers, E. J. (2005). A psychometric analysis of chess expertise. *The American Journal of Psychology, 118*, 29–60.

vanLehn, K., & van der Sande, B. (2009). Acquiring conceptual expertise from modeling: The case of elementary physics. In K. A. Ericsson (Ed.), *The development of professional expertise: Toward measurement of expert performance and design of optimal learning environments* (pp. 356–378). Cambridge, UK: Cambridge University Press.

Wertheimer, M. (1982). *Productive thinking*. Chicago: University of Chicago Press. (Original work published 1945).

Part II
Cognitive Readiness Applications

Chapter 11
Creative Thinking Abilities: Measures for Various Domains

Eunsook Hong

11.1 Introduction

Despite the centuries-long human interest in creativity, it has been only about 60 years since research on creative thinking and creativity received attention by scholars in education. Since Guilford's (1950) historical presidential address on creativity at the American Psychological Association, numerous scholars dedicated their research to exploring and understanding creativity. A number of journals devoted to creativity have sprung up, including *Creativity Research Journal*, *Journal of Creative Behavior,* and *Psychology of Aesthetics, Creativity, and the Arts*. Although there has been a continued effort to conduct scientific studies of creativity as shown in the articles of these journals, research on creativity seems to have been tangential to educational application. The importance of integrating creativity with education emphasized by Guilford (1950), "a comprehensive learning theory must take into account both insight and creative activity" (p. 446), seemed to have been largely ignored by the educational community (Kaufman, Beghetto, Baer, & Ivcevic, 2010; Torrance, 1995).

Even with gifted education, where creativity has been assessed primarily for the identification of gifted and talented students (Kaufman & Baer, 2006), creativity and creative thinking took a back seat, making little impact on the education of the gifted. A recent Newsweek Magazine article, *The Creativity Crisis* (Bronson & Merryman, 2010), described the current status of "American creativity" aptly, indicating that "American creativity is declining." Although studies show that adult accomplishments are related to early creative thinking and creative activities (Hong & Milgram, 2008a; Plucker, Beghetto, & Dow, 2004), schools continue to focus their attention on instruction that requires only analytical-thinking ability and

E. Hong (✉)
Department of Educational Psychology, University of Nevada, Box 453003,
Las Vegas, NV 89154-3003, USA
e-mail: eunsook.hong@unlv.edu

H.F. O'Neil et al. (eds.), *Teaching and Measuring Cognitive Readiness*,
DOI 10.1007/978-1-4614-7579-8_11,

memory-based skills. Creative-thinking ability is either in peripheral vision or out of sight in schools, probably leading to a noticeable decline in recent years (Kim, 2008). It is not very surprising, however, to see this trend in this era of content standards and accountability based primarily on test results, for when the nation is absorbed in a high-stakes testing environment, in which school children must study to prepare for tests for most of the school year, there is simply not much room left for learning that requires reflection or creative thinking.

It is not that there has been no effort to rectify the emphasis on learning environments that deprive learners of opportunities to apply creative thinking. Sporadic efforts to integrate creativity into classroom instruction have been present (Maker, Jo, & Muammar, 2008; Piirto, 2007; Shaheen, 2010; Tomlinson et al., 2002). Sternberg, Grigorenko, and their colleagues (Grigorenko, Jarvin, Tan, & Sternberg, 2008; Sternberg & Grigorenko, 2004; Sternberg, Grigorenko, & Jarvin, 2009) have been steadily making efforts to underscore the importance of creative ability and practical ability beyond analytical ability. A few projects that include measures of creative ability have been tried out in schools and universities, and efforts to instill creativity in the classroom are continuing (Grigorenko & Tan, 2008; Sternberg, 2010). The appearance of an edited book, *Nurturing creativity in the classroom* (Beghetto & Kaufman, 2010), is a hopeful sign indicating that an effort to integrate creativity into classroom instruction is being made in the educational community. Yet, this effort is not in the forefront as educators in general are neither aware of the importance of creativity nor equipped with knowledge and skills to tackle the situation.

Creative-thinking ability is an integral part of cognitive readiness. Individuals, when studying alone or in schools or when working alone or in a team environment, might utilize their knowledge, skills, and competencies related to, for example, problem-solving and decision-making. However, when they encounter unexpected situations during the process of learning and solving problems, mental flexibility is required for appropriate responses to those situations. Creative-thinking ability comes into play when modification of known processes is required or production of original ideas and solutions are preferred or necessary. Creative thinking, as part of cognitive readiness, also signifies that everyday problems can be viewed as requiring creative solutions beyond correct solutions. That is, creative thinking as a disposition, such as openness to new ideas, is a valuable asset for individuals in the twenty-first century.

It is time that faulty conceptions of creativity, such as "people are born creative or uncreative" or "creativity is a fuzzy, soft construct" (Plucker et al., 2004), were debunked and that a climate of understanding that creative thinking and creativity can be developed was generated in schools. In the current global culture, along with the massive use of the Internet by the global community and the uncertainties of the global economy and security, educational systems need to produce individuals who can apply creative thinking as well as analytical thinking. It is time that educational systems showed their capacities to deal with the changes at the global level, met the challenges and demands that the new era has brought to our lives, and generated creative minds.

Recognition of the importance of creativity for advancement of humanity is a prerequisite to cultural and educational support for nurturing creative potential.

The development of expertise requires years of practices (Heller, 2007; Ericsson, 2006, in this book) and support from family and educational systems. Likewise, the development of creative talent requires understanding and support from educational systems and society. Educational systems should invest in the identification and development of creative potential and develop institutional cultures where investment in educating creative minds is recognized as beneficial to society in the long term and to the happiness and wellness of individuals by enabling them to lead a creative lifestyle. We need to develop a culture where creativity and creative thinking is valued highly and acknowledged as a critical resource for human advancement.

11.2 Need for Measures of Creative Potential

Although the role and importance of creativity in individuals' success and the advancement of society have been tacitly acknowledged by the public and educators, misconceptions and stereotypes of creativity persist, largely owing to the inadequacy and inconsistency in the definition and measures of creativity and creative-thinking ability. It is important to distinguish creative thinking from analytical thinking that has been the focus of school learning. Creative-thinking ability is the cognitive ability to generate ideas that are unusual and of high quality, whereas analytical-thinking ability is the cognitive ability to think abstractly and to solve problems logically and systematically. Analytical thinking includes analyzing, comparing, choosing, contrasting, evaluating, judging, inducing, deducing, and inferring (Hong & Milgram, 2008a).

Various theories of creativity have been advanced (Guilford, 1967, 1968; Sternberg & Lubart, 1995; Kaufman & Baer, 2005; Runco, 2007; Simonton, 1997). Although these theories explain the creativity phenomenon well, here I briefly introduce one that sets the stage for this chapter. Hong and Milgram (2008a) proposed two types of manifested talent—*expert talent* and *creative talent*. Either talent (exceptional ability) is developed based on three macro contextual components—cognitive abilities (e.g., analytical and creative-thinking abilities), personal–psychological attributes (e.g., motivation, task commitment), and environmental–social factors (e.g., family, school, mentor). These macro components are foundations for realizing potential talent. An individual who has creative potential might demonstrate good analytical-thinking ability and excellent creative-thinking ability. Individuals with expert talent have highly specialized in-depth knowledge and skills in a particular domain. Individuals with creative talent also have acquired highly specialized knowledge and skills in a domain. However, a difference between the two types of talent is that creative talent generates original and valuable ideas or products, whereas expert talent is evidenced by analytical and problem-solving skills in a particular domain. Creative talent is not required for the demonstration of expertise and expert talent may or may not facilitate creative performance (Hong & Milgram, 2010).

In addition, an individual might be highly motivated and has parents who recognize the potential and are fully supportive. Then this individual might have a good chance to demonstrate creative talent in a domain of his or her interest. The interactive effects of the three contextual components on talent development are complex as the degree of cognitive abilities, various personal traits, and environmental support vary widely.

The development of creative talent is a continuous process as the level and depth of creative accomplishments change from minimal to profound, with a fewer number of people in society who achieve the profound level of accomplishments. Furthermore, the rate of development is determined by the individuals' awareness of their creative potential, by their motivation to seek excellence, by deliberate practice, and with environmental supports. If individuals do not have an interest or intention to act on the opportunities that the environment affords, then learning will not occur. At the same time, when schools and society in general do not afford creativity-fostering climates, individuals' interest in creativity may be suppressed and prematurely recede, thus engendering a situation where the creative spark gets lost.

Although more discussion of various creativity theories should precede the discussion of measurement approaches to creativity, such discussion is beyond the scope of this chapter, as the current chapter focuses on measures of creativity and creative potential. Interested readers can find more information on creativity theories in Kaufman and Sternberg (2010) and Runco (2007).

To help individuals actualize their creative potential (i.e., potential that can be actualized to be highly creative in particular domains), it is important to have quality measures of creative ability. As individuals develop their potential fully. understanding their own level of creative-thinking ability and quality of creative products (i.e., products that are rated as creative by creative talents in the domain) will help them advance to the next level and eventually realize their potential fully. Promoting a creativity-fostering culture should begin with developing quality measures of creative potential, while acknowledging that individuals continue to develop and learn and that while not all individuals exhibit a high level of creative talent, all should have opportunities to develop their potential and to achieve as high a level of talent actualization as they possibly can.

11.3 Defining Creativity and Creativity Thinking

Although creativity is a complex concept and remains somewhat elusive or, some might say, even mysterious, most creativity researchers define creativity as the ability to produce work that is novel (original), appropriate (useful, valuable), and high in quality (Sternberg, Lubart, Kaufman, & Pretz, 2005). Creativity is distinguished from intellectual ability because the products that intelligent people produce may be high in quality and appropriate, but they may not be novel. The *Dictionary of Developmental and Educational Psychology* (Harré & Lamb, 1986) defines "creativity" as the "capacity to produce new ideas, insights, inventions or artistic objects, which are accepted as being of social, spiritual, aesthetic, scientific, or technological value."

"Creative thinking," on the other hand, is defined in the dictionary of the American Psychological Association as mental processes leading with a new invention, solution, or synthesis in any area (Vanderbos, 2006). Dictionary definitions and definitions by individual scholars add some confusion to the distinction among these constructs, with some indicating creativity as a mental process as well as ability (Creativity, 2011) and others using creativity and creative thinking interchangeably as can be found in numerous creativity-related journal articles (Plucker et al., 2004). Nevertheless, the measure of creative-thinking ability is one way to represent creative potential.

One of the reasons creativity is viewed as a fuzzy phenomenon and has inconsistent definitions is due to the vague differentiation between creative potential and creative accomplishment. Creative potential that can be found in children and adolescents may be realized in later years, manifested in their adult creative products and performances. It is partly because creativity researchers and educators sometimes use the term "creativity" when they in fact mean either creative-thinking ability or creative outcomes. This haphazard use of the term "creativity" confuses creative-thinking ability (potential) with creative outcomes or accomplishments. Creative-thinking ability (potential) is a predictor of creative accomplishment, but it is not the outcome per se. Unfortunately, some research studies in the past used creative-thinking ability (e.g., scores from a divergent-thinking test) as a creative outcome of some predictors such as creative personality. Unless the purpose of research is to increase creative-thinking ability as an outcome, creative accomplishments and creative- or divergent-thinking ability should be clearly distinguished. The term and the measures of divergent thinking were originated by Guilford (1950, 1956), and divergent-thinking tests are often used to estimate creative potential.

In this chapter, I attempt to clarify these related but distinct constructs. Plucker et al. (2004) found that more than half of the creativity peer-reviewed journal articles did not define creativity. That is, creativity measured without clear definitions has caused confusion as to which creativity aspects researchers examined in their studies. There are three related constructs intertwined under the name of creativity that may be the main cause for confusion—creativity, creative-thinking ability, and creative products. A definition by Plucker et al. (2004) includes the creativity components of 4 Ps—person, process, product, and press (place or environment)—that have long been discussed since Rhodes (1961); creativity is "the interaction among *aptitude*, *process*, *and environment* by which an individual or group produces a *Perceptible product* that is both *novel* and *useful* as defined within a *social context*" (p. 90).

Combining the definition advanced by Plucker et al. (2004) and a "Comprehensive Model of Giftedness and Talent" forwarded by Hong and Milgram (2008a), I propose a definition of creativity in an effort to distinguish creativity from creative-thinking ability and creative outcomes: "Creativity is a higher-order ability that is manifested in a creative *outcome* (e.g., product, performance, idea, or solution) that is *novel*, *appropriate*, *and of high quality*, resulted from developmental interactions among three broad contextual components—*personal-psychological attributions*, *cognitive abilities*, and *environmental-social factors*." This definition emphasizes the importance of contextual components in the development of creativity as well as the types and standards for creative outcomes.

Creative-thinking ability, a subcomponent of creativity, is the cognitive ability to generate ideas and solutions that are novel and of high quality (Hong & Milgram, 2008a) and is an indicator of creative potential. Creative potential can be developed, and the actualized creative potential is manifested in the form of creative outcomes in particular domains. As indicated earlier, the development of creative talent is a continuous process as the development progresses from minimal to profound. That is, as individuals work toward excellence to their fullest, the level of creative-thinking ability and manifested creative products changes. When speaking of creativity measures, it would be clearer if researchers and practitioners distinguished measures for creative-thinking ability from those of creative products.

It is important to add a note on creative ideas and solutions. The results of creative thinking applied to a problem situation are manifested in creative ideas or creative solutions. They can be viewed as creative outcomes when they are products in the form of ideas and solutions to actual situations, for example, a creative mathematical theorem or a creative idea to rescue Chilean miners trapped underground. On the other hand, when creative-thinking abilities are measured for individuals' ideas and solutions to problems that are often fabricated, these ideas and solutions are considered creative potential, but not creative outcomes.

11.4 Types and Levels of Creativity and Creative Thinking

A number of creativity scholars have classified creativity by its level and quality of creative manifestations. Csikszentmihalyi (1996), Gardner (1994), and Simonton (2000) categorized creativity into Big-*C* creativity and little-*c* creativity, and later Kaufman and his colleagues (2010), included two additional levels to further distinguish the level of creativity. The four-C model of creativity by Kaufman and his associates includes Big-*C*, pro-*c*, little-*c*, and mini-*c* creativity. Big-*C* creativity is related to path-breaking ideas and leads to the transformation of a domain. The Big-*C* creative person is eminent (e.g., Nobel prize winners), a person whose work is well known by people in a particular field. Pro-*c* creativity includes professional-level creators (e.g., award received from a professional association) who have not yet attained a legendary status (Kaufman & Beghetto, 2009). Individuals who exhibit little-*c* creativity use inquisitiveness and imagination in everyday life and link new knowledge to old knowledge, although they have not achieved breakthroughs in their professional domains. The mini-*c* category focuses on the novel and personally meaningful insights and interpretations involved in learning and experience (e.g., a new connection between what has been learned in math and science class). The definition of mini-*c* stresses the importance of personal (or subjective) judgment of novelty and meaningfulness. Mini-*c* insights and interpretations can serve as building blocks on which further creative insights and expression might be produced (Beghetto & Kaufman, 2007; Kaufman et al., 2010). As scholars of creativity ponder what traits and genetic predispositions lead individuals to become Big-*C* creative persons, it is acknowledged that the cultivation of Big-*C* creativity is very difficult (Csikszentmihalyi, 1996; Simonton, 2000; Kaufman et al., 2010). From the perspective of talent development, mini-*c*, little-*c*, and Pro-*c* could be viewed as individuals with creative potentials who can achieve the level of Big-*C*, albeit hard to reach. Although it is interesting and important to understand the Big-*C* individuals and their contributions, this chapter concerns creative potential, specifically in the form of creative-thinking ability.

Other creativity researchers describe the creative-thinking process under the umbrella of creative problem solving (CPS) mostly used in group settings. The decades-old four-stage model by Wallas (1926), the Osborn-Parnes's model (Parnes, 1966, 1988), and the model of Treffinger and his colleagues (Isaksen, Treffinger, & Dorval, 2011; Treffinger, Isaksen, & Dorval, 2003) are representative of CPS models that have been used to increase creative-thinking skills and creative products. Wallas (1926) described creative process as thought processes and actions involved in creative production. This process consists of four stages—preparation, incubation, illumination, and verification—although the process is not linear. The Osborn-Parnes model of CPS is the basis for a program aimed at enhancing the ability of people to develop creative solutions to identified problems. The approach has been widely accepted in business and education (Isaksen & Treffinger, 2004; Isaksen et al., 2011). The CPS process encourages divergent (e.g., brainstorming) as well convergent thinking and includes six stages—mess finding (or objective finding); fact finding; problem finding; idea finding; solution finding; and acceptance finding. Treffinger and his associates extended the Osborn-Parnes model and revised it a few times to reach the current version—*Creative Problem Solving Version 6.1™* (Treffinger et al., 2003). This program guides individuals in using creative- and critical-thinking skills on an individual basis or in a group, helps them understand challenges and opportunities, generates ideas, and develops effective plans for solving problems and managing changes. The CPS Version 6.1™ includes four main components consisting of 8 stages all together: (a) understanding the challenge—constructing opportunities, exploring data, framing problems; (b) generating ideas; (c) preparing for action—developing solutions and building acceptance; and (d) planning your approach—appraising tasks and designing process. Isaksen and Treffinger (2004) reported the history of CPS versions, practice, research, and future research and development needs. Treffinger (2000) contends that the goal of training in CPS is to empower people to use these strategies to deal with the situations and challenges they encounter in real life. As can be seen in these stages, the CPS models, at the macro-process level, can be helpful for planning and executing problem-solving within the prescribed process. However, variations of creative problem-solving processes may be difficult to be accommodated within the CPS models (e.g., Buijs, Smulders, & Van Der Meer, 2009), and the internal cognitive processing involved in the stages of these models is not explained.

Ward, Patterson, and Sifonis (2004) indicated that the way people approach creative idea generation can be varied. This suggests that individuals select cognitive processing strategies when solving creative problems. Selecting effective cognitive strategies in academic problem-solving has demonstrated a positive relationship with performance (e.g., Hong, O'Neil, & Feldon, 2005; Iseman & Naglieri, 2011). Thus, one may predict that this relationship can be evidenced in CPS as well. There has been some research that examined processing variables such as associative basis of creativity (Kasof, 1997; Mednick, 1962) and attentional mechanisms in creative process (Rastogi & Sharma, 2010). However, the scarcity of research on the information processing view of creative thinking is alarming and warrants a multitude of investigations for use in instruction and training (Sawyer et al., 2003; Weisberg, 2006). The recent emergence of neuroimaging studies of creativity (Arden, Chavez, Grazioplene, & Jung, 2010; Dietrich & Kanso, 2010) helps understand brain activities during creative process. However, understanding and the use of neuropsychological information for creativity instruction and training is yet to achieve.

Creative-thinking ability should be one of the important abilities for the twenty-first century that individual learners should acquire and use. The importance of creativity in this century has been emphasized by scholars, especially within the context of the global economy and global technology advancement. As Florida (2007) asserts in "The flight of the creative class: The new global competition for talent," it is imperative that educators be aware of the importance of creative-thinking ability and foster creative thinking in classrooms to produce capable individuals for the twenty-first century.

11.5 Domain Generality and Domain Specificity of Creative-Thinking Ability

The issue of domain generality or domain specificity of creativity has been a topic of research and discussion in the creativity field for the last two decades (Baer, 1994; Hong & Milgram, 2010; Plucker, 1998). The debate has continued partly because creativity has not been defined clearly. Earlier, I distinguished creativity-related constructs into creativity, creative thinking, and creative outcomes. As far as creative outcomes are concerned (i.e., actual creative products, performances, solutions, and ideas), the domain issue should not be much of a question because creative outcomes are manifested in certain domains such as music, science, dance, mathematics, or physics. However, whether creative-thinking ability is domain-general or domain-specific warrants much discussion.

First of all, it is important to summarize the empirical research and thoughts shared by creativity researchers regarding the domain issue. Creative thinking was long considered domain-general (or domain-independent). That is, individuals who score high on a test of creative-thinking ability would be able to generate divergent or original ideas in many domains (Hong & Milgram, 2010). However, scholars have also presented evidence of domain- or task-specificity of creativity (Baer, 1994, 1996, 1998, 2003; Silvia, Kaufman, & Pretz, 2009). A number of studies examined the validity of domain-general ideational fluency-based indices of creative thinking (e.g., producing many ideas, say, all the ways to use a box) as predictors of creative behavior in the real world (Cramond, Matthews-Morgan, Bandalos, & Zuo, 2005; Hong & Milgram, 2010). Some researchers have provided evidence for domain generality referring to the predictive power of domain-general creative-thinking scores on subsequent creative achievement (e.g., Cramond, 1994; Cramond et al., 2005; Torrance, 1993) or to the validity of rating scales of general creative-thinking ability (Batey, Chamorro-Premuzic, & Furnham, 2010; Plucker, 1999; Plucker, Runco, & Lim, 2006). Other investigators viewed the mental process of creative thinking as domain-specific and contended that each domain requires different theoretical and operational definitions (Kaufman & Baer, 2005; Baer, 1998). Some researchers have maintained that creative-thinking ability is both domain-general and domain-specific (Kaufman & Baer, 2006; Plucker & Zabelina, 2009). Regrettably, these studies did not distinguish creative-thinking ability and creative products in their studies of domain specificity/generality of creativity, still leaving some ambiguity on the domain issue, requiring more studies.

Most recently, Hong and Milgram (2010), based on three studies involving six different samples, concluded that domain-general and domain-specific "creative-thinking ability" are distinguishable yet related, and the former had a direct effect on the latter. Interestingly, age/grade level, gender, ethnicity, and learning disability status discriminated general and specific creative-thinking abilities, with the latter ability only being influenced by the demographic variables. Hong and Milgram demonstrated that different life experiences afforded by schooling, gender, and culture may have stronger impacts on domain-specific creative thinking than domain-general creative thinking. That there is a strong relationship between general and specific creative-thinking abilities indicates that although creative ideas and solutions in a specific domain is influenced by experiences and environments, general creative-thinking ability was an important source for individuals' ability to produce creative ideas and solutions in specific domains.

In summary, creativity is a complex concept, including such elements as creative potential, creative outcomes, creative problem-solving process at the individual level and group level, and whether creative-thinking ability is domain-general and domain-specific. Other factors, previously mentioned very briefly, also influence creative achievement (e.g., abilities, traits, social and environmental factors). A bottom line here is that the array of definitions and criteria used to describe creativity has produced various approaches to measuring creative-thinking ability and creative outcomes.

11.6 Measuring Creative-Thinking Abilities

Creativity measures used so far include psychometric tests of divergent thinking (Guilford & Guilford, 1980a, 1980b; Jellen & Urban, 1986; Snyder, Mitchell, Bossomaier, & Pallier, 2004; Torrance, 1974, 1999), inventories that assess creative personality (Cattell, Eber, & Tatsuoka, 1970; Gough, 1979; Torrance & Khatena, 1970), self-efficacy (Beghetto, 2006; Tierney & Farmer, 2011), attitudes (Basadur & Hausdorf, 1996), motivation (Amabile, Hill, Hennessey, & Tighe, 1994), curiosity (Kashdan, Rose, & Fincham, 2004), openness to new experiences (Costa & McCrae, 1991), tolerance of ambiguity (MacDonald, 1970; Norton, 1975), and biographical information (Anastasi & Schaefer, 1969), nominations by peers, teachers, and parents (Pearlman, 1983; Renzulli, Hartman, & Callahan, 1981), product ratings by judges such as consensual assessment technique (Amabile, 1982) or assessment scale (Besemer, 1998; Reis & Renzulli, 1991), self-report of creative activities and accomplishments (Carson, Peterson, & Higgins, 2005; Hocevar, 1979, 1981), creative climate (Amabile, Conti, Coon, Lazenby, & Herron, 1996; Amabile & Gryskiewicz, 1989; Nemiro, 2001), historiometric (Simonton, 1984, 1999), and biographical case-study approaches used to study eminence and creators (Gruber & Wallace, 1999). Recently, more biological methods such as functional magnetic resonance imaging (fMRI), quantitative electroencephalogram (qEEG), or the level of glucose metabolism in the brain measured during the performance of tasks that require creative thinking have been employed to determine the relationship of brain activities with creative thinking (Dietrich & Kanso, 2010; Fink et al., 2009).

In this chapter we focus on measures of creative-thinking abilities—both domain-general and domain-specific. That is, the focus is the measure of creative potential manifested by creative ideas and solutions representing the level of creative-thinking ability, rather than creative outcomes manifested in observable creative products including ideas and solutions as creative outcomes. In addition, since domain-general measures have been widely used in the past several decades, the chapter further concentrates on domain-specific creative-thinking measures. Readers interested in other measurement approaches may refer to Fleenor and Taylor (2004).

Although most creativity scholars agree on the definition of creativity as originality, appropriateness, and quality, the issue of the complexity underlying the creativity concept remains, causing difficulty with operationalizing and developing instruments. Measuring demonstrated talent is not easy, but measuring creative potential, not-yet-realized talents, is even more challenging. Some measures of domain-general creative-thinking ability have been developed and widely used (e.g., Torrance Tests of Creative Thinking). On the contrary, measures of domain-specific creative-thinking abilities are still emerging. In assessing creative-thinking ability, we should be reminded that we are measuring an indicator of creative potential. How and to what extent individuals actually exhibit their realized creative potential depends on the complex interaction of the personal–psychological attributes, cognitive abilities, and social–environmental factors that affect creative-thinking ability and creative outcomes.

11.6.1 Measures of Domain-General Creative-Thinking Ability

Most domain-general tests are based on the work of Guilford (1950, 1956), including the tests developed by Mednick (1962), Meeker, Meeker, and Roid (1985), Milgram and Milgram (1976), Torrance (1999), and Wallach and Kogan (1965). The Torrance Tests of Creative Thinking (TTCT; Torrance, 1974, 1999) has been the most widely used measure. The TTCT was initially published in 1966 and revised in 1999. The TTCT is a standardized, paper-and-pencil measure that is typically group-administered. It consists of two forms—one verbal and one figural. Each form comes in two equivalent versions. The verbal form ("Thinking creatively with words") comprises seven subtests: Asking, Guessing Causes, Guessing Consequences, Product Improvement, Unusual Uses, Unusual Questions, and Just Suppose. The figural form ("Thinking creatively with pictures") consists of three subtests: Picture Construction, Picture Completion, and Parallel Lines/Circles. The cognitive abilities assessed by the TTCT are ideational fluency, flexibility, originality, and elaboration. *Fluency* is based on the number of responses an individual can generate. *Flexibility* is based on the number of different categories an individual's responses can be sorted into. *Originality* refers to the ability to generate uncommon responses. *Elaboration* refers to the ability to elaborate or add details to an idea (Torrance, 1999).

Reasonable reliability and adequate evidence of validity has been reported for the TTCT thus justifying its use in research applications (Cropley, 2000; Kim, 2006; Kim, Cramond, & Bandalos, 2006; Torrance, 2000; Treffinger, 1985). Torrance (1972, 1981, 2002) and Cramond et al. (2005) also demonstrated predictive validity of the TTCT in several longitudinal studies. However, construct validity of TTCT has been questioned in other studies (Almeida, Prieto, Ferrando, & Ferrandiz, 2008; Clapham, 1998; Heausler & Thompson, 1988).

Structure of the Intellect Learning Abilities Test (SOI-LA; Meeker et al., 1985) is based on Guilford's (1956, 1967) Structure of Intellect model of intelligence that differentiated among 180 kinds of thinking, including various forms of divergent thinking. It measures nine cognitive activities connected with creativity, all of them involving divergent thinking applied to three content areas: symbolic, figural, and semantic. Each content area is examined in three subtests (units, relations, and transformations). Each subtest yields scores on fluency, flexibility, originality, and elaboration. Factor analytic studies indicate construct validity and inter-rater reliabilities to be very high (Cropley, 2000; Thompson & Andersson, 1983), although psychometric quality of SOI-LA scores in general has been questioned (Clarizio & Mehens, 1985).

Researchers have used simplified version of creative-thinking (also called divergent-thinking) tests (e.g., Milgram & Milgram, 1976; Wallach & Kogan, 1965). Sample items include "list use of a box" (newspaper, shoe, brick, etc.) or "list all the things you can think of that this figure (a simple line drawing) could represent."

As with all tests, performance on creative-thinking tests is influenced by the examinee's motivational state (Plucker et al., 2006) and the test setting (Runco, Okuda, & Thurston, 1991; Wallach & Kogan, 1965). Even if domain-general tests are not supposed to be influenced by the domain knowledge (i.e., test items are not developed for a certain domain), creative-thinking test questions may ask examinees about things that they may or may not be familiar with. For example, when asked to list things that move on wheels, individuals with experience with various vehicles (toy cars, racing cars) may have an advantage.

11.6.2 Measures of Domain-Specific Creative-Thinking Ability

The measures largely discussed under the banner of domain-generality and domain-specificity research (e.g., Baer & Kaufman, 2005; Silvia et al., 2009) have analyzed creative products rather than creative-thinking ability. Examples of measures in such studies include creating mathematical equations, creating mathematical word problems, writing poetry, writing short stories, and making collages (Baer, 1991, 1993, 1994, 1996, 1998; Han, 2003), musical compositions (Hickey, 2001), everyday problems (Reiter-Palmon, Illies, Cross, Buboltz, & Nimps, 2009), story telling (Han, 2003), and captions written for pictures (Sternberg & Lubart, 1995). These studies used mostly Amabile's (1982, 1996) consensual assessment technique to rate the product or performance. These studies consistently found low or nonsignificant correlations between ratings of creative performance in these different domains or tasks (Baer, 1991, 1993, 1994, 1996; Runco, 1989; Conti, Coon, & Amabile, 1996).

Another type of measure that has been used for creative performance or products is self-report assessment of creative activities and accomplishments (Carson et al., 2005; Hocevar, 1979; Hong & Milgram, 2008b, 2010; Milgram & Hong, 2002). These are domain-specific measures as the instrument specifies the domains in an effort to determine the quantity and quality of activities and accomplishments the respondents engaged in each domain. However, these instruments measure creative performance and products, rather than creative-thinking ability manifested in creative ideas and solutions in specific domains.

Similar to measures of creative performance and products, the development of measures of creative-thinking ability in specific domains requires collaboration between test developers and people with expertise and/or creative talent in particular domains. Unlike the domain-general creative-thinking tests, most of the domain-specific creative-thinking tests are not standardized because the measures are dependent on specific tasks in specific domains that can be as numerous and various as the number and kinds of interests demonstrated by researchers and practitioners for understanding creative thinking in specific domains. This seeming drawback can be rather advantageous to individuals interested in the study of domain-specific creative-thinking ability, as the measure can be adapted to their specific needs. Presented below are some examples of domain-specific measures of creative-thinking abilities.

Creative Real Life Problem Solving (CRLPS) (Casakin, Davidovitch, & Milgram, 2010); Hong & Peng, 2009; Milgram & Hong, 2000–2009) is used to measure creative-thinking ability in various domain- or context-specific problem-solving situations. Each CRLPS measure represents a particular domain/context and consists of two or more items involving real-life problem-solving. Each item describes a problem situation that arises in the specific life situation of prospective respondents in the domain under consideration. When problem-solving items are being developed, potential test respondents' life situations are described in scenarios that could occur in their lives. Although the extent of the possibility of the occurrence of situations depicted in the scenario can vary, each respondent should be able to imagine himself or herself in the scenario while solving the problem. The CRLPS can be used for various ages and for various situations such as classrooms, households, personal lives, or workplaces. Participants are asked to generate as many solutions as they can to each real-life problem presented.

As the CRLPS allows flexibility for various contexts and respondent ages, the level of domain knowledge required to solve the problems in the CRLPS environment is flexible as well. As can be seen in the sample items below, the CRLPS items for young children assume that children have domain knowledge from life experiences afforded in the family, society, and culture they live in. In an architecture example, the respondents who have a higher level of knowledge in the architecture domain would have an advantage in solving the problems more creatively. That is, the level of creative solutions exerted within a given domain/context is determined by creative-thinking ability in the domain as well as the level of domain knowledge. As discussed above, part of this domain-specific creative-thinking ability is influenced by general creative-thinking ability (Hong & Milgram, 2010). Importantly, however, specific creative-thinking ability is influenced also by participants' life experience and domain knowledge.

The importance of domain knowledge in creative accomplishments has been evidenced (Bilalić, McLeod, & Gobet, 2008; Csikszentmihalyi, 1996), although the negative effect of knowledge has been noted as a possible cause for inflexibility (Schooler & Melcher, 1995; Ward & Sifonis, 1997). Csikszentmihalyi (1996), for example, studied individuals known as highly creative in their society in various domains (e.g., scientists, artists, writers, engineers). He found that the primary characteristic of creative individuals is mastery of domain knowledge or skill. Without domain knowledge, creative-thinking ability alone is not likely to lead to creative products that are viewed as high in quality, valuable, and appropriate.

Researchers or practitioners who want to develop CRLPS items should acquire in-depth understandings of respondents' life situations by interviewing potential participants and developing and validating items. The validation would include item content verification with experts in the domain including domain knowledge experts and potential respondents for contextual precision, and item difficulty, clarity, and readability.

Scoring techniques of CRLPS responses are very similar to those for domain-general creative-thinking tests, including fluency (i.e., the total number of discrete responses), flexibility (i.e., the total number of categories), originality (i.e., statistical originality and/or holistic rating of originality by raters knowledgeable of the domain/context/culture), and elaboration (the level of details in the ideas in the solution). Reponses to each item are to be scored by two or more raters who are experienced or trained for rating creative responses and have good domain knowledge. The inter-rater agreement should be computed. For items with different scores, the raters should discuss to reach an agreement. Internal consistency can be estimated as long as there are two or more items. In Hong and Milgram's work (2010), the internal consistencies of CRLPS in three studies ranged from 0.65 to 0.83.

11.6.2.1 Sample Items

(a) School or Classroom Real-Life Situations

Example 1: Domain = Interpersonal problems encountered by elementary school students.

- "The teacher takes all the children in the class outside to play a ball game. She punishes you for speaking to her disrespectfully by making you stay inside the classroom. What can you do? What are all the things that it is possible to do?"

Example 2: Domain = typical academic problems encountered by high school or college students.

- "John prepared for the final test in basic mathematics. He spent many hours, days, and nights summarizing and reviewing the material that had been taught until he felt ready for the test and could get a high grade on it. On the day of the test when he got the test form, he felt that he could not remember anything at all of what he had learned. What are all things that John can do? Suggest as many solutions/ideas as possible."

(b) Teachers Real-Life Situations

Example 1: Domain = Classroom management problems encountered by elementary teachers.

- "You enter the classroom, but the students do not pay any attention and continue doing whatever they were doing as if you had not entered. How will you cope with this situation? Suggest as many ideas/solutions as possible."

Example 2: Domain = Classroom management problems encountered by secondary teachers.

- Your student Stephanie has been disturbing and making noise since the beginning of the class period. At one point she got up and said, "I am sick and tired of studying, I want to stop studying. What will you do? Suggest as many ideas/solutions as possible."

Researchers in various domains have used test questions that measure creative-thinking ability. A few examples that are similar to CRLPS approaches are described below.

Creative Engineering Design Assessment (Charyton, Jagacinski, & Merrill, 2008) was developed to assess creative-thinking ability in engineering design. It consists of five design problems with five parts each to assess an individual's ability to formulate and express design ideas through sketching, providing descriptions, identifying materials, and identifying problems that the design solves. Instructions are to generate as many designs with at least one design per problem. At least one response should be indicated for each of the five questions for each design.

Participants were evaluated according to their originality (0–10, dull to genius), fluency (number of responses), and flexibility (number of response categories defined as variety of responses or number of category type) for each design problem.

Rastogi and Sharma (2010) used real-life problems to assess creative-thinking ability in Sports and Politics domains. As can be seen in the real-life problems, participants were from India and had background knowledge about sport and political affairs in the country. That is, the problems were developed specifically for people who share similar life experiences and culture. The Sports problem stated, "After the dismal performance of the Indian contingent in Olympics 2004, a Sports Review Committee has been formed at the national level. You have been asked to be a member of the Committee and suggest creative measures to improve the standard of sports in the country." The Relationship Problem stated, "As an initiative toward confidence building measures, the Government of your country has called upon certain social groups to discuss the prevalent tension between India and Pakistan. You have been asked to be a member of one such group and suggest creative measures to enhance the friendly ties between India and Pakistan." Responses were rated by analyzing novelty and appropriateness of the two problems.

Owens Creativity Test (OCT; Owens, 1960) was developed to assess mechanical ideation. It consists of two subtests: Power Source Apparatus test and Applications of Mechanisms test. The Power Source Apparatus test asks test takers to generate as many solutions as possible (divergent thinking) to a series of mechanical problems (domain-specific). Test takers are presented with diagrams of mechanical movements, and are asked to draw various ways of producing each of these movements. This subtest produces two scores: the total number of solution and the number of workable solutions. The second subtest, the Application of Mechanisms test, asks the test takers to generate as many ideas as possible for the use of a series of mechanical devices. Test takers are provided with a picture of a device and are asked to report responses in writing. One score is produced by this subtest—the total number of ideas. The OCT had acceptable levels of reliability and validity when applied to engineers in mechanically related jobs (Owens, 1969).

In summary, the above CRLPS questions and other similar real-life questions that measure creative-thinking abilities in specific domains are useful measurement techniques that are relevant for various purposes. The real-life problems are written for specific contexts or domains, and the problem-solving items can be developed to examine prospective respondents for the purposes of (a) selecting potential employees or identify high-potential employees who can creatively plan, design, solve problems, and sell products, (b) engaging individuals in finding creative solutions for their immediate life situations and designing their future, (c) helping individuals become creative thinkers and creative problem solvers, and the list of uses and benefits may be endless.

I want to stress that creative thinking can be taught and developed, starting with awareness of creative thinking and leading to habitual application of creative thinking. As discussed, domain knowledge and skills are an important part of creative thinking and creative production. Thus, the acquisition of domain knowledge and teaching of content knowledge to the mastery level should be stressed. How content knowledge can be taught while fostering creative thinking at the same time would be a topic that deserves another chapter and is a topic that would bring us full circle back to Guilford's address several decades ago.

11.7 Conclusions and Educational Implications

Societies that do not try hard to ensure that the potential talents of individuals are utilized are losing their most valuable natural resource—human capital. Hong and Milgram (2008a) listed a few sources of talent loss. One of those sources regards lack of instruments or lack of use of proper instruments to assess different types and levels of talent. Various instruments are used to measure analytical-thinking ability and memory-based skills in achievement and aptitude tests, but the instruments measuring creative-thinking ability are still evolving. The creative real-life problem-solving approach may be one way to increase the application of tools to increase creative thinking in life in various domains. This type of instrument can be developed with relative ease by content experts working in collaboration with creativity researchers or practitioners. Continued efforts should be made to determine the validity of CRLPS scores, including factorial validity and discriminant validity (Hong & Milgram, 2010) and predictive validity with outcome measures such as creative products and performance.

Human advancement has been accomplished by creators throughout the centuries. Although there may be only a small number of Big-*C* eminent individuals, there can be more individuals with Pro-*c* creativity. Further, there can be many more little-*c* creative individuals in society who can contribute to society in various ways. Individuals can improve creative ability, skills, and dispositions. As an important component of the cognitive readiness required for everyday living as well as for highly time-sensitive situations, creative thinking should be viewed not only as an essential trait but as a trait that needs to be encouraged, stimulated, and educated for further development. Schools should strive to help generate many mini-*c* creative students, finding their creative potentials and supporting them in their development as individuals who produce creative ideas, solutions, or products in domains of their interest. Families, classrooms, and societies and individuals should strive to be aware of their own and others' creative abilities and to create environments that foster creative potential instead of stifling it and accept creative attitudes and creative expression as valuable, as we have long recognized the intellectually capable as important.

Michael Hirsh of Newsweek (2010, August) in his article, "We're No. 11! America may be declining, but don't despair," reported that China now has a much higher rate of patent creation than the United States, as well as a higher number of engineers being graduated. As the world we live in becomes more complex and the problems we face become more challenging, the need for creative individuals, with little-*c* to Big-*C*, to address these situations increases. It is hoped that recognition of the importance of creativity along with an awareness of the declining creative-thinking ability in America (Bronson & Merryman, 2010) may be helpful for turning over a new societal leaf. If a society does not recognize the importance of creativity, it may just head down the very path of "decline" that we have been concerned about, losing the important contributions creative individuals can make. One way to improve the climate at the larger society level is to examine how educational resources are allocated. For example, if most resources associated with research are focused on topics that have to do with testing and academic performance, there will be little left for research on creativity.

Hopefully, over time individuals around the world will become socially, culturally, psychologically, and politically ready to embrace creativity as one of the important factors for human advancement. We, as individuals and society, need both an awareness of importance of creativity and opportunities to practice creativity. The awareness will help us get ready to explore and experience creative insights and activities, and the practice will help us experience creativity unfolding in our daily lives, in classrooms, and in workplaces, so that creativity does not remain merely an abstract notion. Hopefully, the measures of creative potential discussed in this chapter will enable researchers and practitioners help individuals learn to be more creative.

References

Almeida, L. S., Prieto, L. R., Ferrando, M., & Ferrandiz, C. (2008). Torrance Test of Creative Thinking: The question of its construct validity. *Thinking Skills and Creativity, 3*, 53–58.

Amabile, T. M. (1982). Social psychology of creativity: A consensual assessment technique. *Journal of Personality and Social Psychology, 43*, 997–1013.

Amabile, T. M. (1996). *Creativity in context: Update to the social psychology of creativity*. Boulder, CO: Westview.

Amabile, T. M., Conti, R., Coon, H., Lazenby, J., & Herron, M. (1996). Assessing the work environment for creativity. *Academy of Management Journal, 19*, 1154–1184.

Amabile, T. M., & Gryskiewicz, N. (1989). The creativity environment scales: The work environment inventory. *Creativity Research Journal, 2*, 231–254.

Amabile, T. M., Hill, K. G., Hennessey, B. A., & Tighe, E. (1994). The work preference inventory: Assessing intrinsic and extrinsic motivational orientations. *Journal of Personality and Social Psychology, 66*, 950–967.

Anastasi, A., & Schaefer, C. E. (1969). Biographical correlates of artistic and literary creativity in adolescent girls. *Journal of Applied Psychology, 53*, 267–273.

Arden, R., Chavez, R. S., Grazioplene, R., & Jung, R. E. (2010). Neuroimaging creativity: A psychometric view. *Behavioural Brain Research, 214*, 143–156.

Baer, J. (1991). Generality of creativity across performance domains. *Creativity Research Journal, 4*, 23–39.

Baer, J. (1993). *Creativity and divergent thinking: A task specific approach*. Hillsdale, NJ: Erlbaum.

Baer, J. (1994). Divergent thinking is not a general trait: A multi-domain training experiment. *Creativity Research Journal, 7*, 35–46.

Baer, J. (1996). The effects of task-specific divergent-thinking training. *Journal of Creative Behavior, 30*, 183–187.

Baer, J. (1998). The case for domain specificity of creativity. *Creativity Research Journal, 11*, 173–177.

Baer, J. (2003). Evaluative thinking, creativity, and task specificity: Separating wheat from chaff is not the same as finding needles in haystacks. In M. A. Runco (Ed.), *Critical creative processes* (pp. 129–151). Cresskill, NJ: Hampton Press.

Baer, J., & Kaufman, J. C. (2005). Bridging generality and specificity: The Amusement Park Theoretical (APT) model of creativity. *Roeper Review, 27*, 158–163.
Basadur, M., & Hausdorf, P. A. (1996). Measuring divergent thinking attitudes related to creative problem solving and innovation management. *Creativity Research Journal, 9*, 21–32.
Batey, M., Chamorro-Premuzic, T., & Furnham, A. (2010). Individual differences in ideational behavior: Can the big five and psychometric intelligence predict creativity scores? *Creativity Research Journal, 22*, 90–97.
Beghetto, R. A. (2006). Creative self-efficacy: Correlates in middle and secondary students. *Creativity Research Journal, 18*, 447–457.
Beghetto, R. A., & Kaufman, J. C. (2007). Toward a broader conception of creativity: A case for mini-c creativity. *Psychology of Aesthetics, Creativity, and the Arts, 1*, 73–79.
Beghetto, R. A., & Kaufman, J. C. (Eds.). (2010). *Nurturing creativity in the classroom*. New York: Cambridge University Press.
Besemer, S. P. (1998). Creative produce analysis matrix: Testing the model structure and a comparison among products—three novel chairs. *Creativity Research Journal, 11*, 333–346.
Bilalić, M., McLeod, P., & Gobet, F. (2008). Inflexibility of experts—reality or myth? Quantifying the Einstellung effect in chess master. *Cognitive Psychology, 56*, 73–102.
Bronson, P. O., & Merryman, A. (2010, July 19). The creativity crisis. *Newsweek,* 44–50.
Buijs, J., Smulders, F., & Van Der Meer, H. (2009). Toward a more realistic creative problem solving approach. *Creativity and Innovation Management, 18*, 286–298.
Carson, S., Peterson, J. B., & Higgins, D. M. (2005). Reliability, validity, and factor structure of the creative achievement questionnaire. *Creativity Research Journal, 17*, 37–50.
Casakin, H., Davidovitch, N., & Milgram, R. M. (2010). Creative thinking as a predictor of creative problem solving in architectural design students. *Psychology of Aesthetics, Creativity, and the Arts, 4*, 31–35.
Cattell, R. B., Eber, H. W., & Tatsuoka, M. M. (1970). *Handbook for the sixteen personality factor questionnaire (16 PF)*. Champaign, IL: Institute for Personality and Ability Testing.
Charyton, C., Jagacinski, R. J., & Merrill, J. A. (2008). CEDA: A research instrument for creative engineering design assessment. *Psychology of Aesthetics and Creativity in the Arts, 2*, 147–154.
Clapham, M. M. (1998). Structure of figural forms A and B of the Torrance Tests of Creative Thinking. *Educational and Psychological Measurement, 58*, 275–283.
Clarizio, H. F., & Mehens, W. A. (1985). Psychometric limitations of Guilford's structure of the intellect model for identification and programming for the gifted. *Gifted Child Quarterly, 29*, 113–120.
Conti, R., Coon, H., & Amabile, T. M. (1996). Evidence to support the componential model of creativity: Secondary analyses of three studies. *Creativity Research Journal, 9*, 385–389.
Costa, P. T., & McCrae, R. R. (1991). *NEO five-factor inventory*. Odessa, FL: Psychological Assessment Resources.
Cramond, B. (1994). The Torrance Tests of Creative Thinking: From design through establishment of predictive validity. In R. F. Subotnik & K. D. Arnold (Eds.), *Beyond Terman: Contemporary longitudinal studies of giftedness and talent* (pp. 229–254). Norwood, NJ: Ablex.
Cramond, B., Matthews-Morgan, J., Bandalos, D., & Zuo, L. (2005). The Torrance Tests of Creative Thinking: Alive and well in the new millennium. *Gifted Child Quarterly, 49*, 283–291.
Creativity. (2011). In *Dictionary com*. Retrieved from http://dictionary.reference.com/browse/creativity
Cropley, J. (2000). Defining and measuring creativity: Are creativity tests worth using? *Roeper Review, 23*, 72–80.
Csikszentmihalyi, M. (1996). *Creativity flow and the psychology of discovery and invention*. New York: Harper Perennial.
Dietrich, A., & Kanso, R. (2010). A review of EEG, ERP and neuroimaging studies of creativity and insight. *Psychological Bulletin, 136*, 822–848.
Ericsson, K. A. (2006). The influence of experience and deliberate practice on the development of superior expert performance. In K. A. Ericsson, N. Charness, P. J. Feltovich, & R. R. Hoffman

(Eds.), *The Cambridge handbook of expertise and expert performance* (pp. 683–703). New York: Cambridge University Press.

Fink, A., Grabner, R. H., Benedek, M., Reishofer, G., Hauswirth, V., Fally, M., et al. (2009). The creative brain: Investigation of brain activity during creative problem solving by means of EEG and fMRI. *Human Brain Mapping, 30*, 734–748.

Fleenor, J. W., & Taylor, S. (2004). The assessment of creativity. In J. C. Thomas (Ed.), *Comprehensive handbook of psychological assessment: Industrial and organizational assessment* (pp. 75–86). New York: Wiley.

Florida, R. L. (2007). *The flight of the creative class: The new global competition for talent*. New York: Harper Collins.

Gardner, H. (1994). *Creating minds: An anatomy of creativity as seen through the lives of Freud, Einstein, Picasso, Stravinsky, Eliot, Graham, and Gandhi*. New York: Basic Books.

Gough, H. G. (1979). A creative personality scale for the Adjective Check List. *Journal of Personality and Social Psychology, 37*, 1398–1405.

Grigorenko, E. L., Jarvin, L., Tan, M., & Sternberg, R. J. (2008). Something new in the garden: Assessing creativity in academic domains. *Psychology Science Quarterly, 50*, 295–307.

Grigorenko, E. L., & Tan, M. (2008). Teaching creativity as a demand-led competency. In O. S. Tan, D. M. McInerney, A. D. Liem, A. G. Ta, & A. G. Tan (Eds.), *What the West can learn from the East: Asian perspectives on the psychology of learning and motivation* (pp. 11–29). Greenwich, CT: Information Age Press (IAP).

Gruber, H. E., & Wallace, D. B. (1999). The case study method and evolving systems approach for understanding unique creative people at work. In R. J. Sternberg (Ed.), *Handbook of creativity* (pp. 93–115). Cambridge, UK: Cambridge University Press.

Guilford, J. P. (1950). Creativity. *American Psychologist, 5*, 444–454.

Guilford, J. P. (1956). The structure of intellect. *Psychological Bulletin, 53*, 267–293.

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

Guilford, J. P. (1968). *Creativity, intelligence, and their educational implications*. San Diego, CA: EDITS/Knapp.

Guilford, J. P., & Guilford, J. S. (1980a). *Manual for the Christiansen–Guilford Fluency tests*. Orange, CA: Sheridan Psychological Services.

Guilford, J. P., & Guilford, J. S. (1980b). *Manual for the consequences test*. Orange, CA: Sheridan Psychological Services.

Han, K. (2003). Domain-specificity of creativity in young children: How quantitative and qualitative data support it. *Journal of Creative Behavior, 37*, 117–142.

Harré, R., & Lamb, R. (1986). *The dictionary of developmental and educational psychology*. Cambridge, MA: MIT Press.

Heausler, N. L., & Thompson, B. (1988). Structure of the Torrance Tests of creative thinking. *Educational and Psychological Measurement, 48*, 463–468.

Heller, K. A. (2007). Scientific ability and creativity. *High Ability Studies, 18*, 209–234.

Hickey, M. (2001). An application of Amabile's consensual assessment technique for rating the creativity of children's musical compositions. *Journal of Research in Music Education, 49*, 234–244.

Hirsh, M. (2010, August 23 and 30). We're No. 11! America may be declining, but don't despair. *Newsweek*, 40–42.

Hocevar, D. (1979, April). *The development of the creative behavior inventory*. Paper presented at the Annual Meeting of the Rocky Mountain Psychological Association (ERIC Document Reproduction Service No. ED170350).

Hocevar, D. (1981). Measurement of creativity: Review and critique. *Journal of Personality Assessment, 5*, 450–464.

Hong, E., & Milgram, R. M. (2008a). *Preventing talent loss*. New York: Routledge.

Hong, E., & Milgram, R. M. (2008b, 2009). *The activities and accomplishments inventory*. University of Nevada, Las Vegas, NV, and Tel Aviv University, School of Education, Ramat Aviv, Israel.

Hong, E., & Milgram, R. M. (2010). Creative thinking ability: Domain generality and specificity. *Creativity Research Journal, 22*, 272–287.

Hong, E., O'Neil, H. F., & Feldon, D. (2005). Gender effects on mathematics achievement: Mediating role of state and trait self-regulation. In A. M. Gallagher & J. C. Kaufman (Eds.), *Gender differences in mathematics* (pp. 264–293). New York: Cambridge University Press.
Hong, E., & Peng, Y. (2009). *Creative real life problem solving: Thinking and imagination*. Las Vegas, NV: University of Nevada.
Isaksen, S. G., & Treffinger, D. J. (2004). Celebrating 50 years of reflective practice: Versions of creative problem solving. *Journal of Creative Behavior, 38*, 75–101.
Isaksen, S. G., Treffinger, D. J., & Dorval, K. B. (2011). *Creative approaches to problem solving: A framework for innovation and change*. Thousand Oaks, CA: Sage.
Iseman, J. S., & Naglieri, J. A. (2011). A cognitive strategy instruction to improve math calculation for children with ADHD and LD: A randomized controlled study. *Journal of Learning Disabilities, 44*, 184–195.
Jellen, H. G., & Urban, K. K. (1986). The TCT-DP (Test for Creative Thinking-Drawing Production): An instrument that can be applied to most age and ability groups. *Creative Child and Adult Quarterly, 11*, 131–155.
Kashdan, T. B., Rose, P., & Fincham, F. D. (2004). Curiosity and exploration: Facilitating positive subjective experiences and personal growth opportunities. *Journal of Personality Assessment, 82*, 291–305.
Kasof, J. (1997). Creativity and breadth of attention. *Creativity Research Journal, 10*, 303–315.
Kaufman, J. C., & Baer, J. (Eds.). (2005). *Creativity across domains: Faces of the muse*. Mahwah, NJ: Erlbaum.
Kaufman, J. C., & Baer, J. (2006). Intelligent testing with Torrance. *Creativity Research Journal, 18*, 99–102.
Kaufman, J. C., & Beghetto, R. A. (2009). Beyond Big and Little: The four C model of creativity. *Review of General Psychology, 13*, 1–12.
Kaufman, J. C., Beghetto, R. A., Baer, J., & Ivcevic, Z. (2010). Creativity polymathy: What Benjamin Franklin can teach your kindergartener. *Learning and Individual Differences, 20*, 358–364.
Kaufman, J. C., & Sternberg, R. J. (Eds.). (2010). *The Cambridge handbook of creativity*. New York, NY: Cambridge University Press.
Kim, K. H. (2006). Can we trust creativity tests? A review of the Torrance Tests of Creative Thinking (TTCT). *Creativity Research Journal, 18*, 3–14.
Kim, K. H. (2008). Meta-analyses of the relationship of creative achievement to both IQ and divergent thinking test scores. *Journal of Creative Behavior, 42*, 106–130.
Kim, K. H., Cramond, B., & Bandalos, D. L. (2006). The latent structure and measurement invariance of scores on the Torrance Tests of Creative Thinking–Figural. *Educational and Psychological Measurement, 66*, 459–477.
MacDonald, A. P. (1970). Revised scale for ambiguity tolerance: Reliability and validity. *Psychological Reports, 26*, 791–798.
Maker, C. J., Jo, S., & Muammar, O. M. (2008). Development of creativity: The influence of varying levels of implementation of the DISCOVER curriculum model, a non-traditional pedagogical approach. *Learning and Individual Differences, 18*, 402–417.
Mednick, S. A. (1962). The associative basis of the creative process. *Psychological Review, 69*, 220–232.
Meeker, M. N., Meeker, R. J., & Roid, G. H. (1985). *Structure of Intellect Learning Abilities Test (SOI-LA) manual*. Los Angeles: Western Psychological Services.
Milgram, R. M., & Hong, E. (2000–2009). *Ariel real-life problem-solving*. Ariel University Center of Samaria, Ariel, Israel and University of Nevada, Las Vegas, NV.
Milgram, R. M., & Hong, E. (2002). *The activities and accomplishments inventory*. Tel Aviv University, School of Education, Ramat Aviv, Israel, and University of Nevada, Las Vegas, NV.
Milgram, R. M., & Milgram, N. A. (1976). *Tel Aviv Creativity Test (TACT)*. Ramat Aviv, Israel: Tel Aviv University, School of Education.
Nemiro, J. E. (2001). Assessing the climate for creativity in virtual teams. In M. M. Beyerlein, D. A. Johnson, & S. T. Beyerlein (Eds.), *Eighth annual University of North Texas symposium on individual, team, and organizational effectiveness* (pp. 59–84). Oxford: Elsevier Science/JAI Press.

Norton, R. W. (1975). Measurement of ambiguity tolerance. *Journal of Personality Assessment, 39*, 607–612.

Owens, W. A. (1960). *The Owens creativity test*. Ames, IA: Iowa State University Press.

Owens, W. A. (1969). Cognitive, non-cognitive, and environmental correlates of mechanical ingenuity. *Journal of Applied Psychology, 53*, 199–208.

Parnes, S. J. (1966). *Manual for institutes and programs*. Buffalo, NY: Creative Education Foundation Press.

Parnes, S. J. (1988). *Visioning*. Buffalo, NY: Creative Education Foundation Press.

Pearlman, C. (1983). Teachers as an informational resource in identifying and rating student creativity. *Education, 103*, 215–222.

Piirto, J. (2007). *Talented children and adults: Their development and education* (3rd ed.). Waco, TX: Prufrock Press.

Plucker, J. A. (1998). Point counterpoint, beware of simple conclusions: The case for content generality of creativity. *Creativity Research Journal, 11*, 179–182.

Plucker, J. A. (1999). Reanalysis of student responses to creativity checklists: Evidence of content generality. *Journal of Creative Behavior, 33*, 136–137.

Plucker, J. A., Beghetto, R. A., & Dow, G. T. (2004). Why isn't creativity more important to educational psychologists? Potentials, pitfalls, and future directions in creativity research. *Educational Psychologist, 39*, 83–97.

Plucker, J., Runco, M. A., & Lim, W. (2006). Predicting ideational behavior from divergent thinking and discretionary time on task. *Creativity Research Journal, 18*, 55–63.

Plucker, J. A., & Zabelina, D. (2009). And interdisciplinarity: One creativity or many creativities? *ZDM: The International Journal on Mathematics Education, 41*, 5–11.

Rastogi, D., & Sharma, N. K. (2010). Creativity under concurrent and sequential task conditions. *Creativity Research Journal, 22*, 139–150.

Reis, S. M., & Renzulli, J. S. (1991). The assessment of creative products in programs for gifted and talented students. *Gifted Child Quarterly, 35*, 128–134.

Reiter-Palmon, R., Illies, M. Y., Cross, L. K., Buboltz, C., & Nimps, T. (2009). Creativity and domain specificity: The effect of task type on multiple indexes of creative problem-solving. *Psychology of Aesthetics, Creativity, and the Arts, 3*, 73–80.

Renzulli, J. S., Hartman, R. K., & Callahan, C. M. (1981). Teacher identification of superior students. In W. B. Barbe & J. S. Renzulli (Eds.), *Psychology and education of the gifted* (3rd ed., pp. 151–156). New York: Irvinton.

Rhodes, M. (1961). An analysis of creativity. *Phi Delta Kappan, 42*, 305–310.

Runco, M. A. (1989). The creativity of children's art. *Child Study Journal, 19*, 177–190.

Runco, M. A. (2007). *Creativity: Theories and themes: Research, development, and practice*. New York: Elsevier.

Runco, M. A., Okuda, S. M., & Thurston, B. J. (1991). Environmental cues and divergent thinking. In M. A. Runco (Ed.), *Divergent thinking* (pp. 79–85). Norwood, NJ: Ablex.

Sawyer, R. K., John-Steiner, V., Moran, S., Sternberg, R. J., Feldman, D. H., Nakamura, J., et al. (2003). *Creativity and development*. New York: Oxford University Press.

Schooler, J. W., & Melcher, J. (1995). The ineffability of insight. In S. M. Smith, T. B. Ward, & R. A. Finke (Eds.), *The creative cognition approach* (pp. 97–133). Cambridge, MA: MIT Press.

Shaheen, R. (2010). Creativity and education. *Creative Education, 1*(3), 166–169.

Silvia, P. J., Kaufman, J. C., & Pretz, J. E. (2009). Is creativity domain specific? Latent class models of creative accomplishments and creative self-descriptions. *Psychology of Aesthetics, Creativity, and the Arts, 3*, 139–148.

Simonton, D. K. (1984). *Genius, creativity and leadership: Historiometric inquiries*. Cambridge, MA: Harvard University Press.

Simonton, D. K. (1997). Creative productivity: A predictive and explanatory model of career trajectories and landmarks. *Psychological Review, 104*, 66–89.

Simonton, D. K. (1999). Creativity and genius. In L. A. Pervin & O. John (Eds.), *Handbook of personality theory and research* (2nd ed., pp. 629–652). New York: Guilford Press.

Simonton, D. K. (2000). Creativity: Cognitive, personal, developmental, and social aspects. *American Psychologist, 55*, 151–158.
Snyder, A., Mitchell, J., Bossomaier, T., & Pallier, G. (2004). The creativity quotient: An objective scoring of ideational fluency. *Creativity Research Journal, 16*, 415–420.
Sternberg, R. J. (2010). Assessment of gifted students for identification purposes: New techniques for a new millennium. *Learning and Individual Differences, 20*, 327–336.
Sternberg, R. J., & Grigorenko, E. L. (2004). Successful intelligence in the classroom. *Theory Into Practice, 43*, 274–280.
Sternberg, R. J., Grigorenko, E. L., & Jarvin, L. (2009). *Teaching for wisdom, intelligence, creativity, and success*. Thousand Oaks, CA: Corwin.
Sternberg, R. J., & Lubart, T. I. (1995). *Defying the crowd: Cultivating creativity in a culture of conformity*. New York: Free Press.
Sternberg, R. J., Lubart, T. I., Kaufman, J. C., & Pretz, J. E. (2005). Creativity. In K. J. Holyoak & R. G. Morrison (Eds.), *Cambridge handbook of thinking and reasoning* (pp. 351–370). Cambridge, England: Cambridge University Press.
Thompson, B., & Andersson, B. V. (1983). Construct validity of the divergent production subtests from the Structure-of-Intellect Learning Abilities Test. *Educational and Psychological Measurement, 43*, 651–655.
Tierney, P., & Farmer, S. M. (2011). Creative self-efficacy development and creative performance over time. *Journal of Applied Psychology, 96*(2), 277–293.
Tomlinson, C. A., Kaplan, S. N., Renzulli, J. S., Purcell, J. H., Leppien, J. H., & Burns, D. E. (2002). *The parallel curriculum: A design to develop high potential and challenge high-ability learners*. Thousand Oaks, CA: Corwin Press.
Torrance, E. P. (1972). Predictive validity of the Torrance Tests of Creative Thinking. *Journal of Creative Behavior, 6*, 236–252.
Torrance, E. P. (1974). *The Torrance Tests of Creative Thinking*. Bensenville, IL: Scholastic Test Services.
Torrance, E. P. (1981). Predicting the creativity of elementary school children (1958–80) and the teacher who "made a difference". *Gifted Child Quarterly, 25*, 55–62.
Torrance, E. P. (1993). The Beyonder in a thirty-year longitudinal study. *Roeper Review, 15*, 131–135.
Torrance, E. P. (1995). Insights about creativity: Questioned, rejected, ridiculed, ignored. *Educational Psychology Review, 7*, 313–322.
Torrance, E. P. (1999). *Torrance test of creative thinking: Norms and technical manual*. Beaconville, IL: Scholastic Testing Services.
Torrance, E. P. (2000). *Research review for the Torrance Tests of Creative Thinking Figural and Verbal Forms A and B*. Bensenville, IL: Scholastic Testing Service.
Torrance, E. P. (2002). *The manifesto: A guide to developing a creative career*. Westport, CT: Ablex.
Torrance, E. P., & Khatena, J. (1970). What kind of person are you? *Gifted Child Quarterly, 14*, 71–75.
Treffinger, D. J. (1985). Review of the Torrance tests for creative thinking. In J. Mitchell (Ed.), *Ninth mental measurements yearbook* (pp. 1633–1634). Lincoln, NV: Buros Institute of Mental Measurement.
Treffinger, D. J. (2000). *Practice problems for creative problem solving* (3rd ed.). Waco, TX: Prufrock.
Treffinger D. J., Isaksen, S. G., & Dorval, B. (2003). *Creative problem solving* (*CPS version 6.1*^{TM})*: A contemporary framework for managing change*. Sarasota, FL: Center for Creative Learning and Creative Problem Solving Groups.
Vanderbos, G. R. (2006). *APA dictionary of psychology*. Washington, DC: American Psychological Association.
Wallach, M. A., & Kogan, N. (1965). *Modes of thinking in young children: A study of the creativity-intelligence distinction*. New York: Holt, Rinehart, & Winston.
Wallas, G. (1926). *The art of thought*. New York: Harcourt.
Ward, T. B., Patterson, M. J., & Sifonis, C. M. (2004). The role of specificity and abstraction in creative idea generation. *Memory and Cognition, 16*, 1–9.

Chapter 12
Using Analogies as a Basis for Teaching Cognitive Readiness

Keith J. Holyoak and Lindsey E. Richland

An adaptive problem solver needs to be capable of rapidly analyzing novel situations and planning a course of action that is likely to be effective in achieving goals (see Mayer, this book, for additional information on adaptive problem solving). Cognitive readiness, broadly construed, is the state of being prepared to apply one's prior knowledge in an adaptive manner. In this chapter we will review what is known about the role of *analogical reasoning* in fostering cognitive readiness. Analogical reasoning is the process of identifying goal-relevant similarities between a familiar *source* analog and a novel, less understood *target*, and then using the correspondences between the two analogs to generate plausible inferences about the latter (Holyoak, 2012).

The source may be a single concrete situation (e.g., the naval blockade of Cuba in 1962, imposed by the United States in response to placement of nuclear missiles there), a set of multiple cases (e.g., other specific examples of naval blockades), or a more abstract schema (e.g., naval blockades in general). The target may be a situation from a relatively similar domain (e.g., uranium enrichment in Iran, which might be countered by a naval blockade) or a situation in a remote domain (e.g., a computer site launching a cyber attack, which might be countered by an electronic "blockade" designed to cut the threatening site off from the Internet). Cognitive readiness implies that the learner is prepared to transfer knowledge from the

K.J. Holyoak (✉)
Department of Psychology, University of California, Los Angeles, 405 Hilgard Avenue,
Los Angeles, CA 90095-1563, USA
e-mail: holyoak@lifesci.ucla.edu

L.E. Richland
Department of Comparative Human Development, University of Chicago,
5730 S. Woodlawn Avenue, Chicago, IL 60637, USA
e-mail: Lrichland@uchicago.edu

H.F. O'Neil et al. (eds.), *Teaching and Measuring Cognitive Readiness*,
DOI 10.1007/978-1-4614-7579-8_12,

source(s) to a target when the context has been altered, a substantial delay has ensued, and/or the solution applied in the source requires substantial modification to be applicable in the target.

12.1 The Role of Causal Models in Analogical Transfer

Cognitive readiness, and hence successful transfer of knowledge, ultimately depends on the learner's understanding of the source domain. Based on analyses of the use of analogy to support hypotheses in areas of science and the law, Bartha (2010) argues that the reasoner's understanding of the source places an upper bound on the support that the source can provide for a hypothesis about the target. For empirical knowledge (roughly, knowledge about how the world works), understanding can be viewed as the acquisition of a *causal model* of the source domain (Holyoak, Lee, & Lu, 2010; Lee & Holyoak, 2008; Waldmann & Holyoak, 1992). Intuitively, transfer depends on evaluating whether the source and target are alike in terms of the factors that enter into cause-effect relations that impact on goal attainment. For example, the naval blockade of Cuba proved effective in part because Cuba is an island, making it especially dependent on sea transport. If this source analog were used to support a proposal for a naval blockade of Iran, the analogical argument would be weakened by the fact that Iraq is attached to the mainland (as well as by other causally relevant factors, such as Iraq's relatively large size). In contrast, the analogical argument would not be seriously weakened by various other differences between Cuba and Iraq (e.g., the former has a communist government, the latter is an Islamic theocracy) that lack any causal connection to the effectiveness of a naval blockade.

It has long been recognized that understanding the causal structure of the source is critical in producing analogical inferences that enable solving a problem in a novel target. Holyoak (1985) emphasized the centrality of pragmatic constraints on analogical inference that operate in service of goal attainment during problem solving: "the goal is a *reason* for the solution plan; the resources *enable* it; the constraints *prevent* alternative plans; and the outcome is the *result* of executing the solution plan" (p. 70). A well-understood source analog can provide detailed information about a specific pattern of causal events critical to goal attainment, which can be used to provide guidance in making inferences about a complex and poorly understood target analog (see Holyoak & Thagard, 1995).

A series of experiments reported by Lee and Holyoak (2008) demonstrated how causal knowledge guides analogical inference, and that analogical inference is not solely determined by quality of the overall mapping between source and target. Using a common-effect structure (multiple possible causes of a single effect; see Waldmann & Holyoak, 1992), Lee and Holyoak manipulated structural correspondences between the source and the target as well as the causal polarity (generative or preventive) of multiple causes present in the target. In Fig. 12.1, the three panels show examples of causal structures used in their experiments. In the source, three causes (two generative, G_1 and G_2, and one preventive, P) are simultaneously

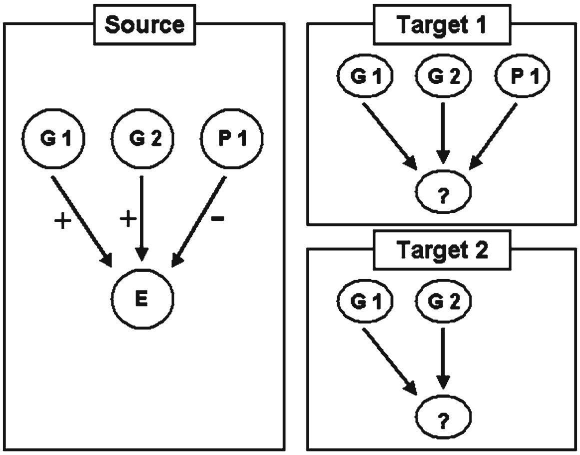


Fig. 12.1 Example of use of causal models in analogical inference. G, P, and E represent a generative cause, preventive cause, and effect, respectively (from Lee & Holyoak, 2008; reprinted by permission)

present, and when the influences of these three causes are combined, the effect occurs. Target analog 1 (G_1G_2P condition) shares all three causal factors with the source, whereas target 2 (G_1G_2 condition) shares only the two generative factors with the source, not the preventive one. Accordingly, target 1 has greater overlap with the source than does target 2, where "overlap" can be defined in terms of some measure of semantic similarity of elements and/or structural matches based on corresponding relations (e.g., Gentner, 1983). All previous computational models of analogical inference (e.g., Falkenhainer, Forbus, & Gentner, 1989; Holyoak & Thagard, 1989; Hummel & Holyoak, 1997, 2003), which assume that the plausibility of target inferences increases monotonically with the overall overlap between the source and target analogs, therefore predict that target 1 is more likely than target 2 to have the effect E.

If analogical inference is guided by causal models, however, the prediction reverses, because dropping a preventive cause, as in target 2 relative to target 1, yields a causal model of the target in which the probability that the effect occurs will *increase*. Lee and Holyoak (2008) tested these alternative predictions using both within-domain analogs (examples of a type of imaginary animal) and cross-domain analogs (imaginary systems in chemistry and astronomy). For both types of materials, they found that people in fact rated analogs in the form of target 2 (G_1G_2) as more likely to exhibit the effect than analogs in the form of target 1 (G_1G_2P; see Fig. 12.2, bottom panel), even though participants rated the target in the G_1G_2 condition as less similar than that in the G_1G_2P condition to the source analog (Fig. 12.2, top panel). Because analogical inferences are based on shared causal structure, such inferences are partially dissociable from overall similarity of the analogs. These findings suggest that understanding human use of analogy to make inferences requires deeper consideration of how causal knowledge is integrated with structural mapping. The results obtained by Lee and Holyoak (2008) imply that people have a natural propensity to focus on causal structure when they use analogies as a source of inferences about an unfamiliar target situation.

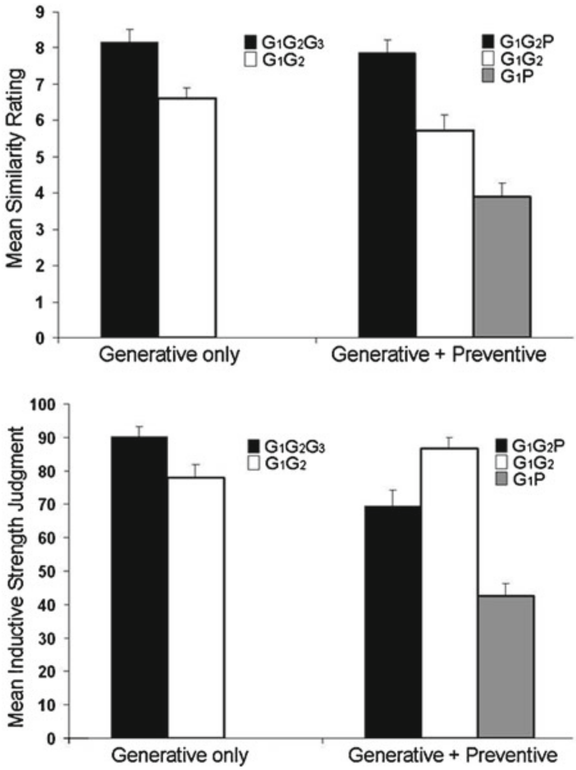


Fig. 12.2 Mean similarity ratings (*top*) and mean inductive strength judgments (*bottom*) for each argument type in the generative-only and generative + preventive conditions of Lee and Holyoak (2008, Experiment 1). *Error bars* represent 1 SEM (from Lee & Holyoak, 2008; reprinted by permission)

Holyoak et al. (2010) developed a computational model of how analogical inference is guided by causal models, formalized within the framework of Bayesian inference. The conceptual framework is schematized in Fig. 12.3. The key assumption of the model is that analogical reasoning uses causal knowledge of the source to develop a causal model of the target, which can in turn be used to derive a variety of inferences about the values of variables in the target. The first stage of analogical inference is learning a causal model of the source (step 1 in Fig. 12.3). The source model is then mapped to the initial (typically impoverished) representation of the target. Based on the mapping (step 2), the causal structure and strengths associated with the source are transferred to the target (Step 3), creating or extending the causal model of the latter. The model of the target can then be "run" (step 4), using causal reasoning to derive inferences about the values of endogenous variables in the target.

For simple causal networks based on binary variables (i.e., cause and effect factors can be either present or absent), Holyoak et al.'s (2010) theory of analogical inference adopts the independently supported assumptions of the Bayesian extension of the power PC theory ("power theory of the probabilistic contrast model"; Cheng, 1997; Lu, Yuille, Liljeholm, Cheng, & Holyoak, 2008). The power PC theory assumes that people use a normative process to integrate the influences of multiple potential causes that co-occur. Holyoak et al. show that their Bayesian model

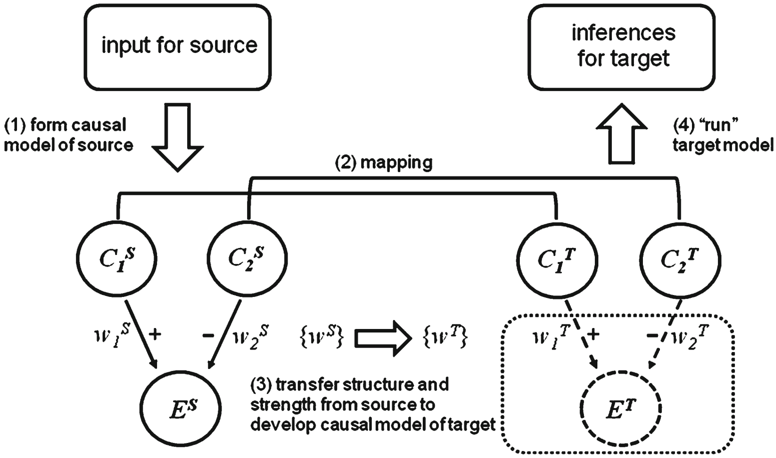


Fig. 12.3 Framework for analogical transfer based on causal models. *Dotted lines* indicate knowledge transferred from source to target (see text) (from Holyoak et al., 2010; reprinted by permission)

provides a close qualitative fit to human judgments for a variety of causal inferences about a target analog based on transfer from a source analog.

The causal-model approach to analogical inference highlights some basic constraints that have implications for cognitive readiness. In particular, the quality of analogical inferences is ultimately limited by the reasoner's causal understanding of the source (Bartha, 2010). If the reasoner fails to understand (or misunderstands) the causal structure of the source, analogical inferences will inevitably suffer. At the same time, transfer is also limited by the quality of the mapping between the source and target. When the correspondences between the analogs are unclear, transfer will also be limited. Critically, not all correspondences are equally important. Rather, correspondences between elements causally related to a reasoning goal are central (Holyoak, 1985; Spellman & Holyoak, 1996). We will now consider some implications of these constraints on analogical transfer for training techniques that can foster readiness to use analogies effectively.

12.2 Analogical Comparison as a Learning Mechanism

Given that analogical transfer is ultimately limited by the learner's understanding of the source domain, it follows that readiness for transfer depends on training techniques that enhance causal understanding of the source. It turns out that one of the most effective techniques for fostering understanding of the source domain is to provide multiple examples and encourage analogical reasoning about the relationships between them. By a kind of "analogical bootstrapping," analogical reasoning can lead to the induction of a more abstract schema, which in turn enables more adaptive transfer.

In an early study of the role of analogical comparison in learning, Gick and Holyoak (1983) employed analogous "convergence" problems as training and transfer tasks. Convergence problems are based on Duncker's (1945) "radiation problem," in which the reasoner must find a way to destroy a stomach tumor without destroying surrounding healthy tissue, by using a type of ray that at sufficiently high intensity will destroy the tumor, but at that same intensity would also destroy the surrounding tissue through which the ray must pass to reach the tumor. The convergence solution requires directing multiple, converging rays toward the tumor at different angles, with the intensity of each ray being sufficiently low to avoid destruction of the surrounding tissue, but the combined intensity of the rays being sufficiently high to destroy the tumor (Gick & Holyoak, 1980).

Gick and Holyoak (1983) trained college students on either one or two convergence problem analogs and tested for transfer with a second analog. An example of an analog for the radiation problem involves a general who wants to amass his forces to attack a fortress, but all the roads leading to the fortress contain mines that will detonate if a sufficiently large group traverses the road. A second analog might involve a fire-fighting scenario in which multiple small hoses are used to extinguish a centrally located fire. In the two-analog condition, participants were asked to compare the two source analogs and write down their commonalities. Gick and Holyoak found that those participants who compared two source analogs exhibited substantially more transfer than did those who saw just one source analog. Those participants who explicitly stated the key aspects of the convergence solution (small forces applied simultaneously from multiple directions) in describing the similarities between the analogs were especially likely to show spontaneous transfer to the tumor problem. Transfer was further enhanced when the concrete analogs were accompanied by a representation of the general solution principle (either as a verbal statement or as a simple diagram). Interestingly, neither the verbal statement nor the diagram conveyed any additional benefit when paired with a single analog.

Gick and Holyoak (1983) argued that encouraging comparison of multiple source analogs (particularly when accompanied by a representation of the solution principle) fostered abstraction of a generalized schema for the problem category, improving the likelihood that subjects will spontaneously recognize the structural similarities among the problems and thereby facilitating transfer (see also Ross & Kennedy, 1990). The critical contribution of analogical comparison to learning is that it highlights the commonalities between two analogs, focusing attention on key aspects causally related to goal attainment. For complex problem situations, the search space of features and relations that may be relevant to a solution will be extremely large. By making an analogical comparison between two examples that mainly share key structural relations, while differing in more surface aspects, the learner will be led to focus on structural relations that might otherwise have been overlooked. The benefit of learning by comparison can be further increased by providing more than two source analogs (Catrambone & Holyoak, 1989).

Very similar results have been obtained in studies in which business students were trained in negotiation strategies (Gentner, Loewenstein, & Thompson, 2003; Loewenstein, Thompson, & Gentner, 1999; Thompson, Gentner, & Loewenstein, 2000). These investigators found that instructions to compare two cases and write down their commonalities (instructions very similar to those used by Gick & Holyoak, 1983) resulted in substantial transfer when the students later engaged in an actual face-to-face bargaining session. Moreover, the benefit of comparison was about three times greater than that provided by presenting two cases (even on the same page) without comparison instructions (Gentner et al., 2003; Thompson et al., 2000; see also Kurtz, Miao, & Gentner, 2001). As in the Gick and Holyoak (1983) study, the quality of the principles elicited during the comparison predicted subsequent transfer.

For much simpler problems, the benefit of comparison as a learning strategy has been demonstrated in children as young as 3 years old (Loewenstein & Gentner, 2001). An important refinement of the use of comparison as a training technique is to provide a series of comparisons ordered "easy to hard," where the early pairs share salient surface similarities as well as less salient relational matches and the later pairs share only relational matches. This "progressive alignment" strategy serves to promote a kind of analogical bootstrapping, using salient similarities to aid the learner in identifying appropriate mappings between objects that also correspond with respect to their relational roles (Kotovsky & Gentner, 1996).

In general, the fundamental benefit of analogical comparison is to foster the generation of a more abstract representation of a class of situations that share important causal structure. When teaching types of problems, a closely related training strategy is to focus students' attention on the explicit subgoal structure of the problems, so that they represent the reasons why actions were performed (Catrambone, 1995, 1996, 1998; Catrambone & Holyoak, 1990). Acquiring explicit knowledge of subgoal structure aids cognitive readiness by making it easier to adapt old solutions from related but non-isomorphic transfer problems (e.g., by deleting, re-ordering, or modifying individual steps in a solution). Teaching subgoal structure has been shown to be especially important in fostering transfer to problems that are relatively dissimilar to the cases used in training (Catrambone, 1998).

12.3 Limits in Processing Resources

The work on learning from comparisons and from cues to subgoal structure demonstrates that simply presenting cases does not suffice to guarantee that learners will be cognitively ready for subsequent transfer to novel problems that share the underlying structure of the training examples. Both developmental studies (Richland, Morrison, & Holyoak, 2006) and neuroimaging studies with young adults (Cho et al., 2010) reveal that analogical mapping is dependent on key executive functions, especially working memory and inhibitory control. Learning from and by analogy thus requires that instructors ensure analogical learning opportunities do not overtax background knowledge, adequate working memory resources, or ability to avoid distraction from surface similarity.

Working memory and inhibitory control are critically involved in two aspects of analogical reasoning: representing and integrating relevant relations and controlling attention to competitive, irrelevant information. Halford and colleagues have hypothesized that processing demand increases as the number of relations to be integrated increases (Halford, 1993; Halford, Wilson, & Phillips, 1998). They have proposed a relational complexity rubric for classifying the processing load of analogies by identifying the number of relations and roles that must be processed in parallel to perform structure mapping (Halford, 1993). These authors used a problem introduced by Sweller (1993) to illustrate the distinction between knowledge and relational complexity: "Suppose five days after the day before yesterday is Friday. What day of the week is tomorrow?" This problem is difficult not because we lack the requisite knowledge of the days of the week, but rather because we must process several of the relationships in parallel.

Attempting to learn from analogies with an overwhelming level of relational complexity reduces learners' readiness and ability to make abstract generalizations (Catrambone, 1995, 1996, 1998; Catrambone & Holyoak, 1990). In the example used above between the 1974 naval blockade of Cuba and the hypothetical "blockade" of a computer cyber attack, the reasoner must hold many key relationships simultaneously in working memory in order to make any sensible inferences. In understanding the source analog, it is important to consider the causal structure of a blockade that involved the threat of nuclear missile placements, the physical configuration of the blockade, proximity between the United States and Cuba, and so forth. Similarly, many relationships must be considered in representing the structure of the target analog, such as the causal structure of the threat posed by the virus, the nature of the computer virus, and the potential number of computers effected. Beyond representing these analogs, the reasoner must then hold the relevant relationships active in order to recognize commonalities and differences (e.g., breadth of blockade necessary, budget constraints, threat potential).

The number of potentially relevant relationships easily overwhelms the working memory system, meaning that cognitive readiness to accomplish this mapping requires the use of tools or strategies for reducing this load. Tools that have been shown to reduce working memory load include providing worked examples (Sweller, 1993, 1994; Sweller, van Merriënboer, & Paas, 1998), introducing visuospatial representations of the key relations (Kosslyn, 1995), breaking the task into subgoals that can be considered separately (Catrambone & Holyoak, 1989), and using gestures to support reasoning (Goldin-Meadow, 2003).

Inhibitory control is also crucial learning from analogy. The ability to control attention and reduce interference from irrelevant but salient features of analogs allows learners to attend to key structural relationships (Cho, Holyoak, & Cannon, 2007; Richland et al., 2006). In the blockade example this might include inhibiting misleading distinctions between the two situations, such as the difference in appearance between a nuclear missile and computer data or between boats and cyber code. While these objects have many differences in features, many of these properties may not be relevant to drawing inferences from one context to another. In addition, there may be irrelevant similarities between these contexts that might lead a reasoner to a false inference or a misconception.

When the demands on inhibitory control are less pronounced, learners have more resources available to allocate to the task of representing and manipulating key similarities and differences between source and target analogs. Designing reasoning contexts that reduce such demands can thus enhance cognitive readiness. Reducing demands on cognitive resources is particularly feasible and important in formulating training contexts. Because novice reasoners have less complete knowledge of the source domain, they are less able to collapse knowledge into groupings, or "chunks," that can reduce processing demands. In addition, when learners are under stress, their inhibitory control resources are reduced, and they are increasingly likely to reason on the basis of object features (Tohill & Holyoak, 2000). Both training and critical problem-solving environments are often highly stressful, in which case strategies for reducing reasoners' attention to irrelevant object properties may be even more important.

12.4 Techniques for Fostering Effective Learning from Training Analogies: Insights from International Peers

Interventions that reduce demands on learners for working memory and inhibitory control during training enhance their readiness to both retain and transfer concepts taught through analogy. A series of studies, described below, have examined feasible, low resource strategies in instructional contexts involving teaching and learning mathematics. Mathematics learning requires abstract, conceptual knowledge that can be transferred across problem contexts. In particular, this series of studies provides insight into the training strategies of peer countries that tend to outperform American mathematics students on international achievement tests (Gonzales et al., 2008).

One robust and somewhat unexpected finding from international studies using achievement tests and videotapes of instructional practices, the technique that forms the basis for this line of research, has been that teaching techniques seem to be culturally organized. While teachers in the United States differ in some expected ways, different American teachers tend to use very similar practices when compared with teachers from other countries (Hiebert et al., 2003; Stigler & Hiebert, 1999). These similarities imply that teaching is a part of culture, suggesting that changes may be difficult as incorporating new training techniques into current practices takes concerted effort. These findings also mean that studying a relatively small sample of teachers internationally is a valid way to gain insight into cultural norms of training practices.

12.4.1 International Practices in Use of Analogy

Richland, Zur, and Holyoak (2007; see also Richland, Holyoak, & Stigler, 2004) studied a subset of videotaped mathematics lessons collected as part of the Trends in International Mathematics and Science Study (Hiebert et al., 2003). The TIMSS 1999 video study used a random probability sampling method to collect videos of everyday eighth-grade classroom teaching in seven countries internationally that outperform the US students in mathematics and or science. Lessons for study were sampled from all classrooms taught in the country to capture variation across the school year and urban, rural, public, private, and religious levels. Richland et al. (2007) took a random subset of these lessons (ten) from the United States and two higher-achieving Asian regions that both outperform the USA regularly, but also are very different from each other in normative teaching practices: Japan and Hong Kong (China). Japanese lessons tend to center on one or two complex problems that students solve independently and then multiple student solutions are compared to introduce new problem-solving strategies and concepts (Hiebert et al., 2003). By contrast, lessons in Hong Kong tend to follow a model that is more similar to the US lesson structure, with a teacher at the board for much of the lesson who engages the students in working through many smaller problems (Hiebert et al., 2003). All instances of instructional comparisons between source and target analogs were identified in every lesson, and then each of those instances was coded to examine the teacher's strategies for supporting students' readiness to notice and make the relational abstraction from the analogy. Coding was conducted by an international team with native speakers from each country yielding high reliability across all codes.

Codes were developed to reflect teachers' common practices that aligned with the cognitive factors outlined above. These codes sought to capture frequency of instructional decisions that could be expected to encourage learners to draw on prior causal knowledge structures, reduce working memory processing load, and reduce demands on inhibitory control. The codes measured the following practical moves by the teacher: (a) constructing visual and mental imagery, (b) using gestures that moved comparatively between the source and target analogs, (c) spatially aligning written representations of the source and target analogs to highlight structural commonalities, (d) using source analogs likely to have a familiar causal structure to learners, (e) making a visual representation of the source analog visible during comparison with the target, and (f) producing a visual representation of a source analog versus only a verbal one.

The results were clear: Asian teachers used all of these cognitive support strategies reliably more frequently than did American teachers. As shown in Fig. 12.4, some strategies were used frequently, others less often, but the Asian teachers were always more likely to include one or more support strategies with their analogies than were teachers in the United States. These differences in strategy use held in spite of there being no differences in the number of analogies used across countries: Americans teachers used high numbers of analogies, but did not provide the same level of support to aid their students in noticing and learning from the key relational structure of the analogies.

Interestingly, few of the analogies in any country involved a source analog that was likely to be very well known to reasoners. Thus as discussed above, learning was often limited by the students' causal knowledge of the source, as they were acquiring and reasoning about the causal structure of both the source and target analogs simultaneously. This type of learning environment, though seemingly very common across instructional environments in different nations, imposes high cognitive demands. In situations involving learning based on an analogy between a less well-known source and a novel target, providing supports for working memory and inhibitory control resources may be particularly crucial.

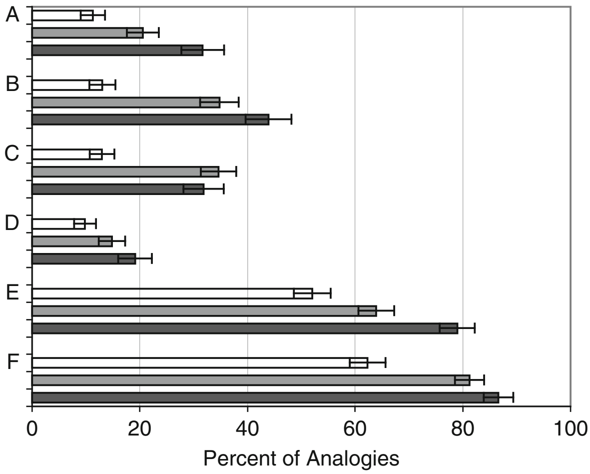


Fig. 12.4 Percent of analogies by region containing cognitive supports: (**a**) visual and mental imagery, (**b**) comparative gesture, (**c**) visual alignment, (**d**) use of a familiar source, (**e**) source visible concurrently with target, (**f**) source presented visually. *White* denotes U.S. teachers, *gray* denotes Chinese teachers, *black* denotes Japanese teachers (from Richland, Zur, & Holyoak, 2007; reprinted by permission)

12.4.2 Experimental Tests of Strategies for Teaching with Analogies

A series of experiments subsequently examined whether using the teaching supports identified in the cross-cultural data described above actually promote learning and transfer when both the source and the target analog are relatively novel. Richland and McDonough (2010) designed instructional videotapes that taught the same two analogs either with or without the support strategies identified in the international video study. Two experiments used materials from the GRE test that were known to be challenging for undergraduates not majoring in a mathematics discipline.

In the first experiment (Richland & McDonough, 2010, Experiment 1), participants were randomly assigned to one of the two conditions: analogy with high cognitive support strategies or analogy with low cognitive support strategies. The content area was permutation and combination problem solving. In both conditions a videotaped teacher first taught and demonstrated a solution to a permutation problem: "Suppose there are five people running in a race. The winner of the race will get a gold medal, the person who comes in second will get a silver medal, and the person who comes in third will get a bronze medal. How many different orders of gold-silver-bronze winners can there be?"

The teacher next taught and demonstrated a solution to a combination problem: "A professor is choosing students to attend a special seminar. She has eleven students to choose from, but she only has four extra tickets available. How many different ways are there to make up the four students chosen to go to the seminar?"

As is evident from these examples, permutation and combination problems share mathematical structure, but there is one key difference. All assigned roles in combination problems are equivalent (e.g., in a combination problem, it does not matter which ticket a student receives). In contrast, order of assignment to roles is critical in permutation problems (e.g., winning gold is different from winning bronze).

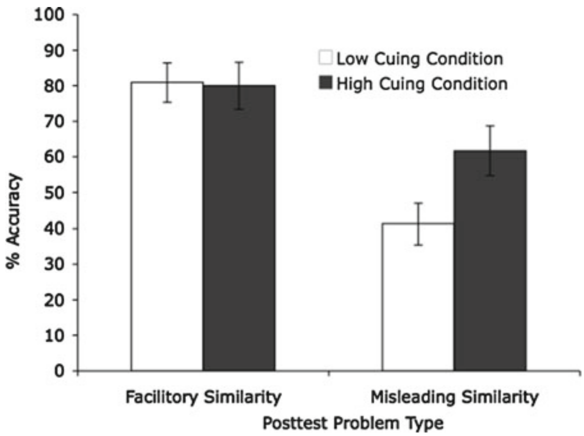


Fig. 12.5 The effects of high versus low cuing of an instructional analogy on posttest problems with varying similarity to instructed problems (from Richland & McDonough, 2010, Experiment 1; reprinted by permission)

In order to solve a combination problem, one divides the total number of permutations by the number of possible role arrangements (in this example problem, 4!, i.e., four factorial, which equals 24 role arrangements).

The posttest was designed to measure both retention of the initial instruction and transfer. The former was assessed using problems with high similarity to the instructed problems, such that permutation problems were set in a race context and combination problems were framed as tickets to a lecture. Transfer was assessed using "misleading similarity" problems in which the surface context and the mathematical structures were cross-mapped; a permutation problem was instantiated as tickets to a lecture, and a combination problem framed as a race. Performance on these problems reflected learners' ability to represent the source and target problem based on mathematical structure as opposed to surface features.

As evident in Fig. 12.5, the data revealed a significant interaction between instructional condition and problem type. There were no differences on the facilitatory similarity problems, revealing that participants in both instructional conditions benefited from the instructional analogy, showing approximately 80 % accuracy on these near-transfer problems (baseline performance with the same population was 7 %). The results were different for the misleading similarity problems (baseline level 10 %). Participants who had received instruction with high support cues were more likely to solve the misleading similarity problems by transferring on the basis of structure versus surface features (62 % accuracy). These participants significantly outperformed participants who had been trained on the low support video, whose accuracy was 41 %. This pattern indicates that either type of instructional analogy was beneficial but that adding supports to aid learners' cognitive readiness led to more adaptable, schematic knowledge representations.

This interaction was replicated in two additional experiments. Richland and McDonough (2010, Experiment 2) obtained a very similar result with undergraduates learning to solve proportion word problems through an analogy between a correct solution and a common but invalid use of the linearity assumption. A third experiment replicated the result in a classroom context with school-age children learning division of rational numbers by analogy to division of natural numbers (Richland & Hansen, in press).

Overall, these data reveal a reliable finding that high quality analogies can be effective tools for enhancing cognitive readiness for learning but that including supplemental strategies to provide cognitive support maximizes the impact of analogy training on transfer. When given such cues, learners seem to have developed more conceptual, schematized representations of the instructed concepts and more adaptive proficiency in representing new problems. These tools require only a small investment on the part of the instructor, yet have the potential for broad learning gains. Importantly, though seemingly quite simple, these strategies for reducing processing load do not appear to be traditional parts of typical training strategies used by teachers in the United States, certainly not in teaching mathematics.

12.5 Conclusions

In general, teaching relational structure constitutes a powerful tool for fostering cognitive readiness for transfer. Instruction based on analogy is not straightforward, however, since limits in relevant knowledge and processing capacity increase the likelihood that learners fail to notice or benefit from analogies in teaching. The aim of the teacher should be to assist the learner in developing veridical causal models of the domain or deep understanding of content structures. Major strategies for using analogies effectively in teaching include guided comparison of examples, highlighting of relations by principles and visual diagrams, ordering examples to encourage progressive alignment, and focusing attention on subgoals.

Other supportive cues can improve acquisition of the underlying relational structure. Useful interventions include reducing processing load, facilitating attention to relational structure of target problems, drawing learners' attention to relations versus object features, reducing competitive interference, and encouraging learners to draw on prior knowledge. The benefits of relational instruction are most apparent when the learner is later faced with novel problems that require extension and adaptation of the earlier examples used in training.

We end with a list of recommendations for practices that can be customized to different learning contexts and training needs. In many training contexts analogies are a widely used, but under-considered, resource for enhancing abstraction and transfer. Comparing a new problem or concept to prior knowledge is a cognitive ability deeply embedded into our thinking and perhaps is what makes human an especially adaptive species (Penn, Holyoak, & Povinelli, 2008). However, explicit training by analogy is not as naturally reliable, and instructional analogies can be greatly improved by using several key support strategies. These tools are maximally important when learners have incomplete knowledge of the domain, are under stress, or are otherwise operating with limited cognitive resources.

12.6 Practical Recommendations for Improving Cognitive Readiness Through Analogy

Training strategies for enhancing cognitive readiness through analogy are collapsed here into a useful list for the reader's reference. Citations and fuller descriptions are available in the text above:

1. Use a source analog with a causal structure that is well known to learners
2. Have learners compare two or more source analogs before transferring to a target
3. Guide learners through comparisons between analogs, either with explicit instructions or mapping tasks
4. Order source analogs from "easy" mappings to more challenging ones, fostering progressive alignment between the analogs
5. Reduce processing load on the working memory system. This can be accomplished in several ways, including:
   (a) Break target problem into subgoals that can be accomplished separately
   (b) Use visual or mental imagery
   (c) Create visual representations of source and target analogs, rather than describing them only verbally
   (d) Make visual representations of source and target analogs visible simultaneously
6. Reduce demands on inhibitory control. This can be accomplished in several ways, including:
   (a) Design source and target visual representations in ways that highlight the key relational correspondences and downplay irrelevant similarities and differences
   (b) Use hand gestures that move between the representations of source and target correspondences to draw attention to relational commonalities

Acknowledgements The work reported herein was partially supported by grants from the Office of Naval Research, Award Numbers N000140810186 and N000140810126. The findings and opinions expressed here do not necessarily reflect the positions or policies of the Office of Naval Research. We thank Harry O'Neil for helpful comments on an earlier draft.

References

Bartha, P. (2010). *By parallel reasoning: The construction and evaluation of analogical arguments*. Oxford, UK: Oxford University Press.

Catrambone, R. (1995). Aiding subgoal learning: Effects on transfer. *Journal of Educational Psychology, 87*, 5–17.

Catrambone, R. (1996). Generalizing solution procedures learned from examples. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 22*, 1020–1031.

Catrambone, R. (1998). The subgoal learning model: Creating better examples so that students can solve novel problems. *Journal of Experimental Psychology: General, 127*, 355–376.

Catrambone, R., & Holyoak, K. J. (1989). Overcoming contextual limitations on problem-solving transfer. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15*, 1147–1156.

Catrambone, R., & Holyoak, K. J. (1990). Learning subgoals and methods for solving probability problems. *Memory & Cognition, 18*, 593–603.

Cheng, P. W. (1997). From covariation to causation: A causal power theory. *Psychological Review, 104*, 367–405.

Cho, S., Holyoak, K. J., & Cannon, T. (2007). Analogical reasoning in working memory: Resources shared among relational integration, interference resolution, and maintenance. *Memory & Cognition, 35*, 1445–1455.

Cho, S., Moody, T. D., Fernandino, L., Mumford, J. A., Poldrack, R. A., Cannon, T. D., et al. (2010). Common and dissociable prefrontal loci associated with component mechanisms of analogical reasoning. *Cerebral Cortex, 20*, 524–533.

Duncker, K. (1945). On problem solving. *Psychological Monographs, 58*(5), 270.

Falkenhainer, B., Forbus, K. D., & Gentner, D. (1989). The structure-mapping engine: Algorithm and examples. *Artificial Intelligence, 41*, 1–63.

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science, 7*, 155–170.

Gentner, D., Loewenstein, J., & Thompson, L. (2003). Learning and transfer: A general role for analogical encoding. *Journal of Educational Psychology, 95*, 393–408.

Gick, M. L., & Holyoak, K. J. (1980). Analogical problem solving. *Cognitive Psychology, 12*, 306–355.

Gick, M. L., & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology, 15*, 1–38.

Goldin-Meadow, S. (2003). *Hearing gesture: How our hands help us think*. Cambridge, MA: Harvard University Press.

Gonzales, P., Williams, T., Jocelyn, L., Roey, S., Kastberg, D., & Brenwald, S. (2008). *Highlights from TIMSS 2007: Mathematics and science achievement of U.S. fourth- and eighth-grade students in an international context* (NCES 2009–001). Washington, DC: National Center for Education Statistics, U.S. Department of Education.

Halford, G. (1993). *Children's understanding. The development of mental models*. Hillsdale, NJ: Erlbaum.

Halford, G. S., Wilson, W. H., & Phillips, S. (1998). Processing capacity defined by relational complexity: Implications for comparative, developmental, and cognitive psychology. *The Behavioral and Brain Sciences, 21*(6), 803–831.

Hiebert, J., Gallimore, R., Garnier, H., Givvin, K. B., Hollingsworth, H., Jacobs, J., et al. (2003). *Teaching mathematics in seven countries: Results from the TIMSS 1999 video study* (NCES 2003–013). Washington, DC: National Center for Education Statistics, U.S. Department of Education.

Holyoak, K. J. (1985). The pragmatics of analogical transfer. In G. H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 19, pp. 59–87). New York: Academic Press.

Holyoak, K. J. (2012). Analogy and relational reasoning. In K. J. Holyoak & R. G. Morrison (Eds.), *The Oxford handbook of thinking and reasoning* (pp. 234–259). New York: Oxford University Press.

Holyoak, K. J., Lee, H. S., & Lu, H. (2010). Analogical and category-based inference: A theoretical integration with Bayesian causal models. *Journal of Experimental Psychology: General, 139*, 702–727.

Holyoak, K. J., & Thagard, P. (1989). Analogical mapping by constraint satisfaction. *Cognitive Science, 13*, 295–355.

Holyoak, K. J., & Thagard, P. (1995). *Mental leaps: Analogy in creative thought*. Cambridge, MA: MIT Press.

Hummel, J. E., & Holyoak, K. J. (1997). Distributed representations of structure: A theory of analogical access and mapping. *Psychological Review, 104*, 427–466.

Hummel, J. E., & Holyoak, K. J. (2003). A symbolic-connectionist theory of relational inference and generalization. *Psychological Review, 110*, 220–263.

Kosslyn, S. M. (1995). *Image and brain: The resolution of the imagery debate*. Cambridge, UK: Cambridge University Press.

Kotovsky, L., & Gentner, D. (1996). Comparison and categorization in the development of relational similarity. *Child Development, 67*, 2797–2822.

Kurtz, K. J., Miao, C., & Gentner, D. (2001). Learning by analogical bootstrapping. *The Journal of the Learning Sciences, 10*, 417–446.

Lee, H. S., & Holyoak, K. J. (2008). The role of causal models in analogical inference. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 34*, 1111–1122.

Loewenstein, J., & Gentner, D. (2001). Spatial mapping in preschoolers: Close comparisons facilitate far mappings. *Journal of Cognition and Development, 2*, 189–219.

Loewenstein, J., Thompson, L., & Gentner, D. (1999). Analogical encoding facilitates knowledge transfer in negotiation. *Psychonomic Bulletin & Review, 6*, 586–597.

Lu, H., Yuille, A. L., Liljeholm, M., Cheng, P. W., & Holyoak, K. J. (2008). Bayesian generic priors for causal learning. *Psychological Review, 115*, 955–982.

Penn, D. C., Holyoak, K. J., & Povinelli, D. J. (2008). Darwin's mistake: Explaining the discontinuity between human and nonhuman minds. *The Behavioral and Brain Sciences, 31*, 109–130.

Richland, L. E., & Hansen, J. (in press). Reducing cognitive load in learning by analogy. *International Journal of Psychological Studies*.

Richland, L. E., Holyoak, K. J., & Stigler, J. W. (2004). The role of analogy in teaching middle-school mathematics. *Cognition and Instruction, 22*, 37–60.

Richland, L. E., & McDonough, I. M. (2010). Learning by analogy: Discriminating between potential analogs. *Contemporary Educational Psychology, 35*, 28–43.

Richland, L. E., Morrison, R. G., & Holyoak, K. J. (2006). Children's development of analogical reasoning: Insights from scene analogy problems. *Journal of Experimental Child Psychology, 94*, 249–271.

Richland, L. E., Zur, O., & Holyoak, K. J. (2007). Cognitive supports for analogies in the mathematics classroom. *Science, 316*, 1128–1129.

Ross, B. H., & Kennedy, P. T. (1990). Generalizing from the use of earlier examples in problem solving. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 16*, 42–55.

Spellman, B. A., & Holyoak, K. J. (1996). Pragmatics in analogical mapping. *Cognitive Psychology, 31*, 307–346.

Stigler, J. W., & Hiebert, J. (1999). *The teaching gap*. New York: The Free Press.

Sweller, J. (1993). Some cognitive processes and their consequences for the organisation and presentation of information. *Australian Journal of Psychology, 45*, 1–8.

Sweller, J. (1994). Cognitive load theory, learning difficulty, and instructional design. *Learning and Instruction, 4*, 295–312.

Sweller, J., van Merriënboer, J. J. G., & Paas, F. G. W. C. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*, 251–296.

Thompson, L., Gentner, D., & Loewenstein, J. (2000). Avoiding missed opportunities in managerial life: Analogical training more powerful than individual case training. *Organization Behavior and Human Decision Processes, 82*, 60–75.

Tohill, J. M., & Holyoak, K. J. (2000). The impact of anxiety on analogical reasoning. *Thinking and Reasoning, 6*, 27–40.

Waldmann, M. R., & Holyoak, K. J. (1992). Predictive and diagnostic learning within causal models: Asymmetries in cue competition. *Journal of Experimental Psychology: General, 121*, 222–236.

Chapter 13
Simulation Assessment of Cognitive Readiness

Klaus Breuer and Siegfried Streufert

13.1 Introduction

Actions that are designed to have an impact on a challenging task environment require readiness to deal with the problems at hand, no matter what they might be. Readiness *including readiness to make decisions* within complex task settings involves several aspects of cognitive functioning, among them (1) motivation, (2) specific content knowledge of the task components, generally generated by training and/or experience, and (3) the capacity to deal effectively with multiple components of the task, their interrelationships and their interplay over time. For the purpose of this chapter, we will specifically focus on the third of these components. Let us assume, for present purposes, that adequate motivation is present and that content knowledge or technical skill of the task is at hand. As a matter of fact, for most highly professional individuals functioning over considerable time in repeatedly challenging (especially managerial and decision making) task settings, these aspects are, at least in good part, given. Still, success in responding to a difficult and/or challenging task is often not attained or insufficient. As a result, resolving a certain particular problem at hand may not be possible. Yet, often the lack of success is not due to the structure of the task itself, but due to inadequate cognitive readiness for dealing with the multiple interrelationships among task components and with their changes over time.

Tasks differ widely in their demands (what kind of task it is and what kind of knowledge, experience and/or technical and cognitive competency may be required).

K. Breuer (✉)
Johannes Gutenberg-Universität, Am Rodelberg 15, 55131 Mainz (JGU), Germany
e-mail: breuer@uni-mainz.de

S. Streufert
Professor Emeritus at Penn State College of Medicine, Hershey, PA, USA
e-mail: sstreufert@streufert.net

H.F. O'Neil et al. (eds.), *Teaching and Measuring Cognitive Readiness*,
DOI 10.1007/978-1-4614-7579-8_13,

Let us distinguish between at least three "types" of tasks (Streufert & Swezey, 1986): those that are *simple*, those that are *complicated*, and those that are *complex*. For both simple and complicated tasks, procedural (content) knowledge and related skills are adequate to achieve the desired outcome. Of course, many simple and complicated tasks exist. Yet various other tasks are "complex." Challenges by complex tasks may be generated when the task environment is volatile, when uncertainty about the setting or the intended outcome prevails, when the interplay of task components reflects complexity, when task requirements are ambiguous, and when feedback about consequences of actions taken to deal with the task are slow (delayed beyond the need to take subsequent actions). Task settings of that kind have been described as VUCAD (volatility, uncertainty, complexity, ambiguity, and delayed feedback) (Satish & Streufert, 2004; Streufert, 1993). The first four among these characteristics were coined by the US military (US Army War College) to describe decision making in military campaigns, and the last was added by complexity theorists.

When we deal with VUCAD, we can no longer expect that we will be able to discover the "perfect" and certainly not the "correct" solution to problems at hand. Readiness now means the capacity to find a good (at least more than adequate) solution via continuous active (re)orientation and continuous adaptation that monitors and adjusts activities to generate and maintain sufficiently effective outcomes over time. Successful dealing with VUCAD requires continued awareness and utilization of multiple task components (challenges); it requires an understanding of possibly changing events and event interrelationships (impacts of events upon one another). It requires monitoring those components, as well as the (potentially interactive) effect of each component upon the intended outcome. Whenever we are concerned with managerial effectiveness in today's world, we need to ask ourselves some pertinent questions: Does an individual who has to make important decisions possess the capacity to function effectively under VUCAD? How certain can we be, that this individual will attain an excellent outcome? And, if the person of interest is not yet cognitively ready when confronted with VUCAD, is training toward greater effectiveness possible? How can it be achieved?

13.2 Model-Based vs. Scenario-Based Approaches

To answer such questions, theorists and researchers have employed different approaches to human readiness under VUCAD task demands. Some theorists and researchers concerned with complex problem solving and decision making have created computer-based models and have exposed decision makers to such settings. Often the model components were derived from theory or from attained "insights" of the theorist. Such approaches are in good part based on systemic models and, to some extent, on anticipated time change effects. Relevant variables are defined in a computer program that *represents* the theoretically specified dynamics of the environment. The features of the model are at least in part hidden from participants who, over time, may discover some of those variable characteristics (including feedback

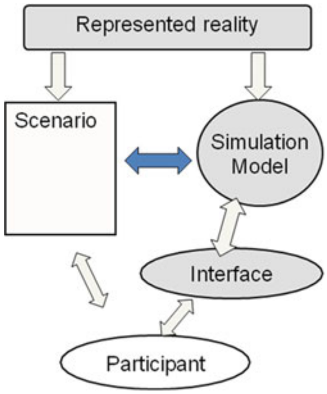


Fig. 13.1 The model-based approach to diagnostic complex simulations

loops and time delays), once feedback is received in response to a decision. In other words, once decisions are made, the decisions have an impact on the task setting (based on the computational model). Subsequent decisions by the participant have to deal with the modified task setting, and so forth (Fig. 13.1).

From a measurement (diagnostic) standpoint, the model status in such designs is partially confounded with the participants' decision-making process (Breuer, Molkenthin, & Tennyson, 2006; Breuer & Streufert, 1995a, 1995b). An example of such an approach is the micro-world methodology developed by Doerner (e.g., 1996; Dörner & Wearing, 1995). In Doerner's designs, the initial characteristics of a task environment, as well as resources (decision options), are presented to participants at the beginning of their task. Participant(s) then make(s) a sequence of decisions. Each set of decisions generates specific (model calculated) outcomes that modify the task requirements prior to the next set of decisions. Outcomes are based both on the interrelations between the modeled variables and on the changes of the variables status over time. The micro-world approach has been frequently utilized to diagnose managerial effectiveness (Funke, 1993; Hussy, 1998; Tennyson & Breuer, 2002). Since the action of the participant(s) as well as the model characteristics affect outcomes, a micro-world is able to demonstrate that specific action patterns lead to failure of even expert participants (e.g., Doerner, Kreuzig, Reither, & Stäudel, 1983). Nonetheless, this methodology has its limits if we wish to measure (diagnose) the capabilities and readiness of an individual who must deal with VUCAD (Molkenthin, Breuer, & Tennyson, 2008).

13.3 Free Simulation Technologies

While micro-worlds differ from most simulations and in-basket techniques by typically requiring sequential interactions (decisions—changed setting response—decisions—response, etc.) between participant and computer, micro-worlds nonetheless have much in common with most other simulation technologies. Fromkin and

Streufert (1976) defined such methods as “free simulations” where actions of the participants have a direct impact on changes in the task environment over time. Free simulations allow the introduction of highly complex scenarios that can challenge decision makers with continuous VUCAD environments. Which decisions are made and whether and how resources are utilized are entirely under the control of decision makers. Participants typically enjoy the experiences, in part because they can encounter the consequences of their own actions over time; they can readjust their actions and modify their approaches to attain desired goals.

Many free simulations have been employed in the military, business, medicine, and in other fields where VUCAD is encountered by decision makers. Yet these methods still suffer from the same measurement problem that we encountered with micro-worlds (cf., Streufert & Swezey, 1985). Because of the “free” nature of the participant–task interaction, precise measurement of performance is severely restricted. Evaluation of performance within free simulations must be left to a (potentially biased or even unreliable—yet mostly well trained) observer who estimates effectiveness based on specific activities and decisions made. Measurement precision is necessarily restricted because different decision makers generate diverse subsequent environments to which they (again) respond in different unique ways. Comparisons among decision makers are therefore restricted. In other words, each participant ends up with a different flow of events. As a consequence, the reliability of performance measurement remains in some question.

13.4 Quasi-Experimental Simulations

An attempt to resolve such problems was made by Streufert and his associates (Streufert & Streufert, 1981), as well as subsequently by Breuer and associates (Breuer & Satish, 2003; Breuer & Streufert, 1995a, 1995b) as well as Satish and associates who developed or utilized “quasi-experimental simulations.” This approach avoids the exposure of participants to the internal dynamics that, for example, characterize the model aspect of micro-worlds. As in other simulation technologies, the quasi-experimental approach exposes the participant(s) to a VUCAD setting. However, in this method, all important (relevant to measurement) events that occur during the simulation are preprogrammed in content and time, i.e., all participants are exposed to the same sets of inputs at identical time points. Some structurally unimportant (to task and to measurement) events can be influenced by actions of the participant decision maker(s) to provide the impression of a responsive task environment (Fig. 13.2).

Because of preprogrammed information inputs (task events) over time, precise performance measurement becomes possible. Using this method, at least two kinds of independent variables can be introduced into research and training methodologies: (1) individual differences (e.g., experience, training, specific competency levels, etc.) and (2) environmental challenge characteristics (e.g., task load, stress, etc.). Furthermore, environmental (VUCAD relevant) task characteristics can be

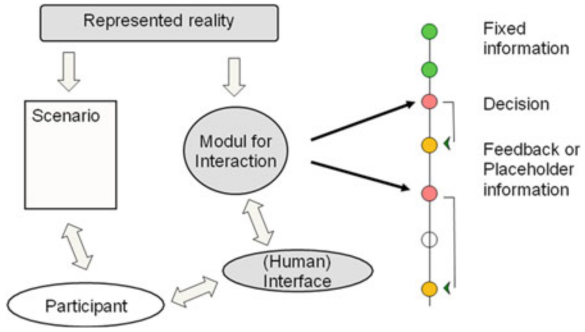


Fig. 13.2 The quasi-experimental approach to diagnostic complex simulations

changed or specifically varied across time when useful. Moreover (3) multiple performance characteristics (among them response frequency, strategic capacity, response to stress, and many others) can be assessed. Finally, (4) because the participants are exposed to identical experiences, comparisons of performance across individuals, across manipulated environmental conditions, and across other introduced variables of interest can be obtained and, most of all, can be validated.

13.5 Complexity and Meta-complexity

Quasi-Experimental technology is scenario-based, not model-based. It provides the participant with a VUCAD environmental setting, resources to deal with events in that setting over time and information (as stated earlier in good part preprogrammed in both content and timing) about events that occur in that setting. Participants are able to make decisions at any time and can make decisions of any kind, as long as the resources to make a particular decision are (or remain) available. Typically participants get deeply involved in the task (high motivation) and their responses mirror (established validity) their behavior in normal real-world tasks. Since the scenarios appear "familiar" to participants (for example, from newscasts or media), but neither mirror the participants' own job characteristics nor their prior experience, the resulting measured behavior indicates the individual's *underlying* capacity to function (cognitive readiness) in response to VUCAD settings.

Early research using quasi-experimental simulations tended to focus on determining whether "cognitive complexity" was evident in participants' actions (e.g., Streufert, 1970). That approach, however, *did not* take into account that task characteristics can differ widely: Some tasks or task components merely require simple procedural action; others are best handled by a breadth of approach that considers choices among two or more alternative solutions; yet others (where considerable VUCAD is present), necessitate multifaceted functioning that has been described both by cognitive (Streufert, 1997) and by science wide (e.g., Kauffman, 1992, 1995, 2002) complexity theory. Cognitive readiness to deal effectively with the task environment must, in part, depend on the specific task at hand, no matter whether it represents a simple procedural task or a multifaceted task involving VUCAD, and so forth. An approach that takes account of diverse task requirements, encompassing simple, intermediate and highly complex functioning, and the appropriate handling of each is described by meta-complexity (e.g., Streufert, 2005). Quality of performance is based on effective functioning that takes account of the individual's optimal handling of the specific task and its characteristics at hand. Contemporary research with quasi-experimental simulation technology takes account of meta-complexity.

13.6 Measuring Decision-Making Effectiveness

The Strategic Management Simulations (SMS) are quasi-experimental technologies that were developed to assess and train *multiple aspects of* decision-making competence *that are discussed in the next paragraph* (Streufert, 1970; Streufert & Swezey, 1986). A number of matched (in task demands and measurement outcome) scenarios have been developed and widely used to measure cognitive readiness across professional specialties (e.g., Shamba, Woodline County, Astaban). Other scenarios have been specifically developed for clients with particular interests.

If we intend to generate an assessment of a person's actual decision-making competence, we need to provide a setting that generates indicators of that competence. For that purpose, two requirements are of necessity: complexity of the task and time to utilize those competencies. The SMS, which provides the basis for this chapter, provides both. Participants are exposed to a highly complex (multifaceted) simulation that dynamically and meaningfully changes over time. Secondly, while the time point where the simulation ends is not stated to the participant beforehand, it continues over six half hour periods of real time while simulated time may reflect days, weeks, or even months (depending on the internal logic of a specific scenario).

The SMS and its predecessors were initially developed by Streufert and associates (e.g., Streufert, 1970; Streufert & Streufert, 1978). To generate an inclusive list of decision-making abilities, these authors collected more than 90 measurement-based indicators of decision making. Data were obtained from several 100 participants across continents and finally subjected to statistical techniques that identified the degrees of overlap or independence of the measurement technologies (such as multidimensional scaling, varimax factor analysis, and so forth). The generated data indicated (again across nations) a set of between 9 and 12 *independent* measures of decision making that go beyond knowledge/experience and motivation, i.e., measures that do assess decision-making competence independently of each other. The most common nine measures are listed in Table 13.1.

Based on subsequent validity data (see below), some of the measures were subsequently subdivided into components that are based on a common overall

Table 13.1 Nine basic measures of decision -making competence/cognitive readiness

#	Label	Definition
1	Activity	The number of decisions made
2	Response speed	Delay between receipt of information and initial decision response
3	Responsiveness	Focus of actions on task at hand
4	Initiative	Uncued actions taken
5	Information orientation	Information search activity
6	Emergency responsiveness	Actions taken in response to received emergency information
7	Breadth	Alternative actions to deal with the task at hand
8	Planning	Reference to future decisions (plans) in present decision
9	Strategy	Effective use of a prior action to facilitate a subsequent action

competence, yet where people who are more successful frequently differ from people who are only moderately successful. For example, in reference to basic measure 9, measures of strategy include

1. Contextual strategy: strategy is used in a specific context.
2. Basic strategy: an assessment of the frequency in which strategy is used overall.
3. Encompassing strategy: strategy is utilized across multiple aspects of the task.
4. Advanced strategy: strategic action interconnect multiple aspects of the task toward common goals.
5. Strategic complexity: multiple sequential strategic coupling of actions over task aspects and over time toward multiple often interrelated goals.

The SMS assess multiple specific decision-making competencies at a general level, focusing on the underlying competence that decision makers would bring to a wide variety of divergent situations. As such, these simulations are useful across professional specializations and across cultures and languages. They have been validated in various business contexts, in pharmacology (e.g., effects of drugs on decision-making competence; cf., Streufert & Gengo, 1993), medicine (Streufert & Satish, 2003), crisis management (Streufert—emergency decision making; Breuer & Satish, 2003), and more. The simulations provide validated measurement of competence (cognitive readiness to handle various levels of tasks) in terms of a set of quantitative scores and in terms of visually effective "qualitative" graphic representations of functioning.

13.7 Quantitative Measurement

As already suggested above, scenario-based (non-model-based) simulations where all events related to measurement are preprogrammed generate information on the performance of individuals that is subject to direct quantitative measurement.

We can determine (meta-complexity, meta-readiness) whether the response to a specific task component is relevant, e.g., whether the participant is sensitive to the particular level of task demands (e.g., handling a simple procedural task in the same fashion one handles a VUCAD task would not be useful). We can determine whether a participant does or does not engage in specific behaviors (actions) that are of interest, how often—and relevant to what kinds of information—those actions do occur, whether those actions are related to other actions as part of overall or specific planning and/or strategy, whether behavior changes in kind (either effectively or ineffectively) occurs in response to stress, to emergencies, to failure experience, and more. All of these (and more) performance characteristics are numerically scored by a computer program, eliminating the problem of observer error or bias. Validity, reliability, and independence (factor structure) of the obtained measures have been repeatedly demonstrated across cultures and languages over several decades (e.g., Streufert, Pogash, & Piasecki, 1988).

13.8 Qualitative (Graphic) Measurement Representation

In addition to quantitative measurement, a qualitative graphic representation of the multiple components of performance can be generated (cf., Breuer & Satish, 2003; Streufert & Satish, 1997). While these graphs are "qualitative" in terms of their visual communication, they are nonetheless based on the same hard "quantitative" data that are considered in the section above. Described as "Time/Event Matrices" these graphs plot time (typically several hours of simulation participation) on the horizontal axis and decision categories on the vertical axis. Each decision made is identified as a point (vertically) above the time where the decision occurs and (horizontally) at the level of the decision category to which it refers to.

If a participant in the simulation makes a decision that is an intended antecedent of a future decision (involving planning and/or strategy), the first decision can be connected with the second (once the second decision is carried out) via a diagonal with an arrowhead pointing toward the second decision (reflecting use of strategy). If the later decision is not carried out (either because other actions took care of the problem or because the decision maker forgot or neglected future action), the arrow becomes a vertical line, pointing to the decision category that was planned (reflecting planning that was not followed up). If a decision maker finds a previous action useful to generate a new decision (but the later decision had not been preplanned), the latter decision is connected to the earlier decision via a diagonal line with the arrowhead pointing toward the earlier decision (reflecting utilization of opportunity). When information received during simulation participation is utilized to generate a particular decision, the point of information receipt is marked with a star (horizontally) ahead of the relevant decision type and vertically above the time point of information receipt. A decision that utilizes received information as (at least part of) the reason for making the decision is circled.

13.9 Reliability and Validity

The SMS reliability is excellent for test-retest (e.g., Streufert et al., 1988): data that have been obtained across the different SMS scenarios show high reliability (0.8–0.94 across different measures). Test-retest results were obtained on 2 subsequent days, a week apart, and for about 30 participants 1 year apart. Meaningful test-retest data can be obtained as long as participants do not know (unless told) what the simulation measures. The different information content of the SMS scenarios, despite the equivalent task demands, prevents participants from learning how to perform better—unless, of course, they become trained (for information on training, see below).

No matter how reliable a measurement technique may be, if it is not valid it is not useful. Although there are a number of validity studies supporting the prediction of success for performance on the various simulation measures (e.g., Breuer & Streufert, 1995a, 1995b; Funke, 1993), one striking example may be sufficient to make the point. It is well known to almost everybody, and it is demonstrated as well as widely accepted in the behavioral sciences that more than a minimum of alcohol consumption has negative effects on human functioning, including upon cognitive capacity (readiness). This frame of reference has been used as an anchor for a series of studies on drug effects on cognitive readiness. In a double-blind placebo-controlled effort, meaning that neither the participants nor the administrators of the simulation runs knew the treatment conditions effective in any one simulation run, decision makers were exposed to placebo (disguised as alcohol), to alcohol exposure at the 0.05 level or at the 0.10 level. Maintenance of blood alcohol was measured by breathalyzers with the data collected by a researcher who was not administering the simulation. The individuals participated in three different SMS scenario runs in randomized order. Measures of cognitive functioning have been assessed across the established profile (compare Table 13.1 and Fig. 13.3). A plot of the respective results is presented in Fig. 13.4 (Streufert & Pogash, 1998). As predicted performance was worse under 0.05 alcohol than under placebo and much worse under treatment that generated the 0.10 alcohol blood level. Performance under alcohol conditions was worse for almost all measures under placebo condition. Performance under 0.1 alcohol level was worse in 21 out of 24 measures compared to the 0.05 level of blood alcohol. Similar results were obtained for treatment with a tranquilizer (Streufert et al., 1996) and with certain other (psychoactive) drugs that are able to cross the blood–brain barrier. Together with findings from additional research, the conclusions for the validity of the simulation measures upon cognitive functioning are substantiated.

13.10 Training

Of course we could train most motivated individuals to be effective in dealing with a procedural task where “right” and “wrong” responses would be identified in advance. But can we train all managers to become qualified decision makers when

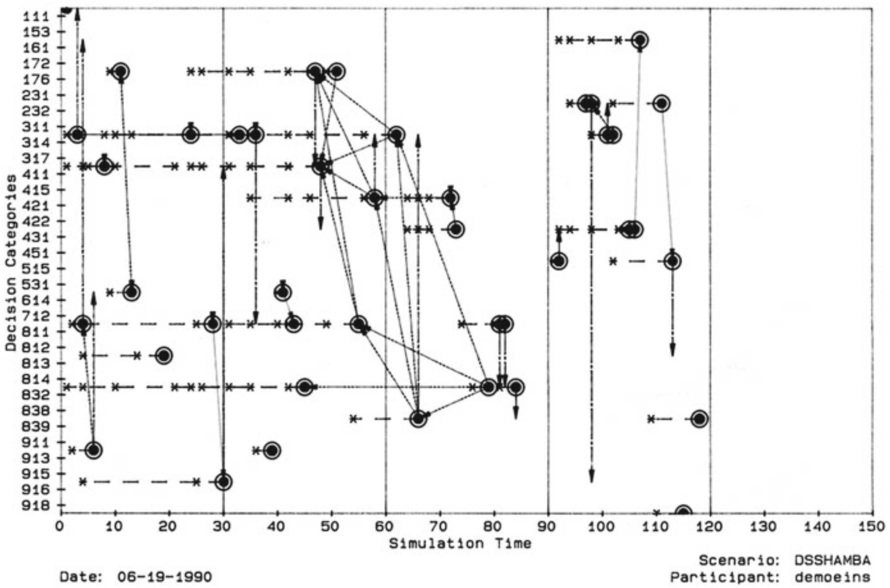


Fig. 13.3 A time/event matrix representing the decision-making process of a participant

VUCAD strikes? What would be the procedure to generate cognitive readiness to function effectively enough under such challenging conditions?

Past research appears to suggest that not everyone (interestingly enough irrelevant of measured intelligence) is trainable to handle VUCAD (Streufert & Streufert, 1978; Streufert & Swezey, 1985). Some basic capacity to deal effectively with VUCAD has to be present. In many cases an individual may be able to function under VUCAD—something that science-wide complexity theorists might call the "edge of chaos" (Kauffman, 1995)—in some (or few) specific task settings. Where that capacity is present in one realm or another, it can be expanded to other task challenges. One should note, however, that training and learning to deal with *very specific* ambiguous, complex, and delayed feedback settings may merely reflect the acquisition of a highly complicated procedure that becomes useless when uncertainty and especially volatility create major changes in task demands.

Effective training toward improved functioning on specific SMS measures has been reported. Interviews with supervisors on the job have also shown that training after simulation participation generates improved functioning. Interestingly enough, training of individuals with prior mild to moderate head injury, utilizing practice with training vignettes developed by Streufert has had some strikingly favorable results. Still, the underlying capacity to deal with VCAD must be present, before training efforts can be generally effective.

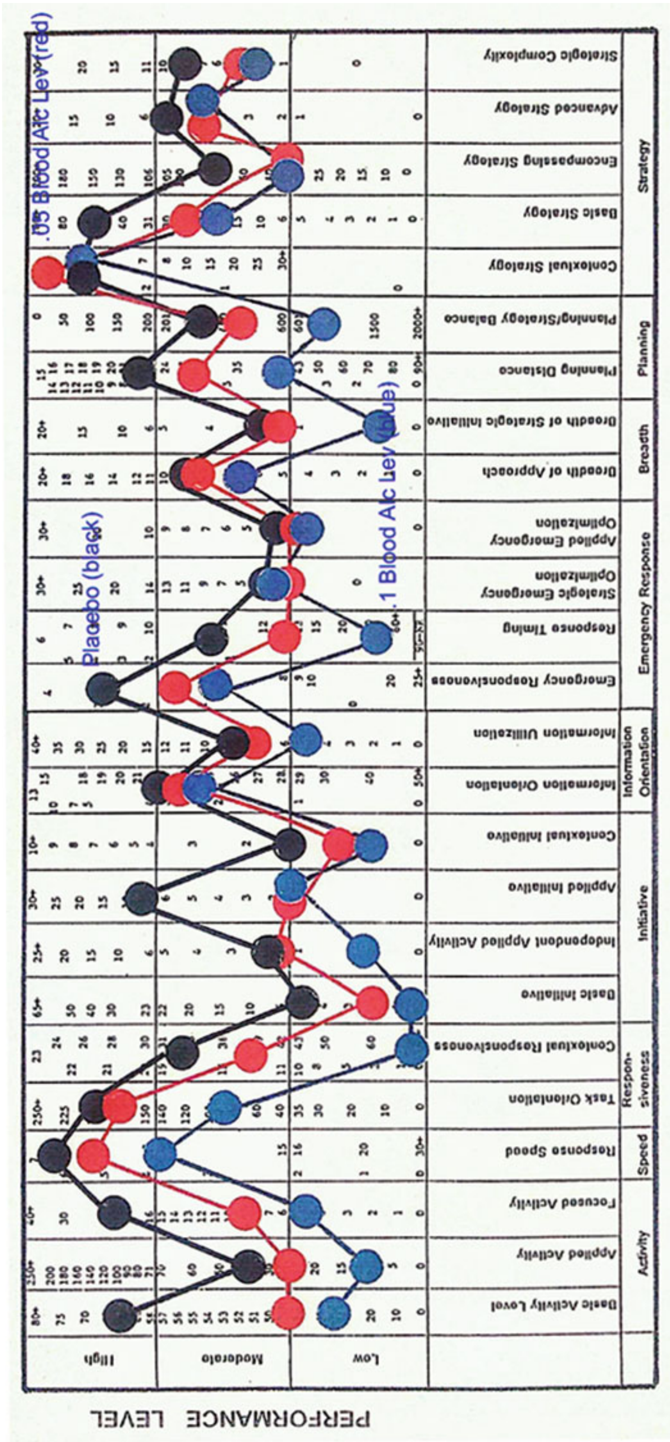


Fig. 13.4 Validation results based on double-blind "treatment" with alcohol

13.11 Future Efforts

Without question, cognitive readiness to perform real-world tasks is of importance if we wish to obtain successful functioning and meaningful productivity. In this chapter we have focused on tasks that involve VUCAD, yet we have recognized that tasks differ. We must consider readiness in terms of the task demands, the existing level of relevant competence of the individual involved and the degree to which differences (or changes) in task demands over time translate into effective performance. General use of the SMS (see above) can certainly generate data we need to select, place, and evaluate individuals whose relevant competence is matched to the task environment. We could select managers that "match" task-specific demands. And finally, we can train individuals to deal better with various levels of demands, including task challenges that involve VUCAD (e.g., Haritz & Breuer, 1995).

Beyond the potential application of SMS as it has been used in the past, specific versions of the quasi-experimental simulation approach, related techniques, and their associated measurement technologies might and probably should be developed to match specific challenges that occur in specific work environment settings (e.g., Breuer et al., 2006; Molkenthin et al., 2008). Such efforts have already been effectively utilized in specific environments in the air force, medicine, business administration, and in some other fields.

Acknowledgements The work reported herein was partially supported by a grant from the Office of Naval Research, Award Number N000140810126. The findings and opinions expressed here do not necessarily reflect the positions or policies of the Office of Naval Research.

References

Breuer, K., Molkenthin, R., & Tennyson, R. D. (2006). Role of simulation in web-based learning. In H. F. O'Neil & R. S. Perez (Eds.), *Web-based learning: Theory, research, and practice* (pp. 307–326). Mahwah, NJ: Erlbaum.

Breuer, K., & Satish, U. (2003). Emergency management simulations: An approach to the assessment of decision-making processes in complex dynamic crisis environments. In J. J. Gonzaléz (Ed.), *From modeling to managing security—A system dynamics approach* (pp. 145–156). Kristiansand: Norwegian Academic.

Breuer, K., & Streufert, S. (1995a). Strategic management simulations the German case. In M. Mulder, W. J. Nijhoff, & R. O. Brinkerhoff (Eds.), *Corporate training for effective performance* (pp. 195–208). Norwell: Kluwer Academic.

Breuer, K., & Streufert, S. (1995b). Computerunterstützte Eignungsdiagnostik mit komplexen dynamischen Szenarios: Ausräumen von Missverständnissen [Computer-based diagnostics of aptitudes by means of complex dynamic scenarios: Clarification of misunderstandings]. *Zeitschrift für Arbeits- und Organisationspsychologie, 39*, 34–36.

Doerner, D. (1996). *The logic of failure*. New York: Metropolitan Books, Henry Holt.

Doerner, D., Kreuzig, H. W., Reither, F., & Stäudel, T. (Eds.). (1983). *Lohhausen—Vom Umgang mit Unbestimmtheit und Komplexität*. Bern: Huber.

Dörner, D., & Wearing, A. J. (1995). Complex problem solving: Towards a theory. In P. A. Frensch & J. Funke (Eds.), *Complex problem solving: The European perspective* (pp. 65–99). Hillsdale, NJ: Erlbaum.

Fromkin, H. L., & Streufert, S. (1976). Laboratory experimentation. In M. D. Dunette (Ed.), *Handbook of industrial and organizational psychology* (pp. 415–465). Chicago: Rand-McNally.

Funke, U. (1993). Computerunterstützte Eignungsdiagnostik mit komplexen dynamischen Szenarios [Computer-based diagnostics of aptitude by means of complex dynamic scenarios]. *Zeitschrift für Arbeits- und Organisationspsychologie, 37*, 109–118.

Haritz, J., & Breuer, K. (1995). Computersimulierte und dynamische Entscheidungssituationen als Element der multikulturellen Personalentwicklung. In J. M. Scholz (Ed.), *Internationales Change-Management [International change management]* (pp. 109–120). Stuttgart: Schäffer-Poeschel.

Hussy, W. (1998). Denken und Problemlösen/Thinking and Problem-Solving (2. extended edition). Stuttgart: Kohlhammer.

Kauffman, S. A. (1992). *Origins of order: Self-organization and selection in evolution*. New York: Oxford University Press.

Kauffman, S. A. (1995). *At home in the universe: The search for laws of self organization and complexity*. New York: Oxford University Press.

Kauffman, S. A. (2002). *Investigations*. New York: Oxford University Press.

Molkenthin, R., Breuer, K., & Tennyson, R. D. (2008). Real-time diagnostics of problem-solving behavior for business simulations. In E. L. Baker, J. Dickieson, W. Wulfeck, & H. F. O'Neil (Eds.), *Assessment of problem solving using simulations* (pp. 201–221). Mahwah, NJ: Erlbaum.

Satish, U., & Streufert, S. (2004). Training to defend: SMS simulations. *ISCRAM 2004 Book of Abstracts, 42*.

Streufert, S. (1970). Complexity and complex decision making: Convergences between differentiation and integration approaches to the prediction of task performance. *Journal of Experimental Social Psychology, 6*, 491–509.

Streufert, S. (1993). Assessment/training of senior personnel with the strategic management simulations. In *35th Conference of the Military Testing Association* (pp. 429–433), Williamsburg, VA.

Streufert, S. (1997). Complexity: An integration of theories. *Journal of Applied Social Psychology, 27*, 2068–2095.

Streufert, S. (2005). Emergency decision making and metacomplexity. In B. Van de Walle & B. Carle (Eds), *Information systems for crisis response and management* (pp. 67–73). SCK.CEN and Universiteit van Tilburg ISCRAM 2005.

Streufert, S., & Gengo, F. M. (1993). *Effects of drugs on human functioning*. Basel, Switzerland: Karger.

Streufert, S., & Pogash, R. (1998). The effect of limited alcohol consumption on complex task performance. In J. Snell & M. M. Lorist (Eds.), *Nicotine, caffeine and social drinking* (pp. 331–346). Amsterdam: Harwood.

Streufert, S., Pogash, R., & Piasecki, M. (1988). Simulation-based assessment of managerial competence: Reliability and validity. *Personnel Psychology, 41*, 537–557.

Streufert, S., & Satish, U. (1997). Graphic representation of processing structure: The time-event matrix. *Journal of Applied Social Psychology, 27*, 2122–2132.

Streufert, S., & Satish, U. (2003). Strategic management simulations: From psychology research to medical practice. *American Psychological Society Observer, 16*, 9–54.

Streufert, S., Satish, U., Pogash, R., Gingrich, D., Landis, J. R., Lonardi, L., et al. (1996). Effects of alprazolam on complex human functioning. *Journal of Applied Social Psychology, 26*, 1912–1930.

Streufert, S., & Streufert, S. C. (1978). *Behavior in the complex environment*. New York: Wiley.

Streufert, S., & Streufert, S. C. (1981). Experimentation in behavioral medicine: In search of an applicable method. In J. E. Eiser (Ed.), *Social psychology and behavioral medicine* (pp. 63–76). New York: Wiley.
Streufert, S., & Swezey, R. W. (1985). Simulation and related research methods in environmental psychology. In J. Singer & A. Baum (Eds.), *Advances in environmental psychology* (Vol. 5, pp. 99–117). Hillsdale, NJ: Erlbaum.
Streufert, S., & Swezey, R. (1986). *Complexity, managers and organizations*. New York: Academic.
Tennyson, R. D., & Breuer, K. (2002). Improving problem solving and creativity through use of complex-dynamic simulations. *Computers in Human Behavior, 18*, 650–668.

Chapter 14
Assessing Cognitive Readiness in a Simulation-Based Training Environment

William L. Bewley, John J. Lee, Barbara Jones, and Hongwen Cai

14.1 Introduction

To begin, we ask the fundamental question posed by this volume: what is cognitive readiness? We have adopted the Fletcher (2004) definition: cognitive readiness is the mental preparation needed to perform competently "in the complex and unpredictable environment of modern military operations" (p. 1). One very good example of such an environment is tactical decision making in the Combat Information Center (CIC) of a U.S. Navy warship, especially the performance of the Tactical Action Officer (TAO). This context has been the focus of our work.

The Navy Surface Warfare Officers School (SWOS) is the schoolhouse responsible for training surface ship TAOs within the Department Head (DH) course. We have been conducting research and developing performance assessments used in the Department Head course for several years.

A Department Head (DH) is an officer in charge of a ship department (e.g., engineering, weapons, combat systems, and operations). The current TAO training at SWOS consists of 12 weeks (out of a total of 27 weeks of DH instruction), covering all warfare areas, with training provided through a mix of computer-based testing (CBT) and extensive self-study, rigorous instructor-led classroom training, multiple written exams to assess level of knowledge, and tactical scenarios and practicals to apply their knowledge. Most important is the use of the Multi-Mission Tactical Trainer (MMTT) to practice, demonstrate, and assess skills, including building the

W.L. Bewley (✉) • J.J. Lee • B. Jones
National Center for Research on Evaluation, Standards, and Student Testing (CRESST)/UCLA, Los Angeles, CA, USA
e-mail: bewley@cse.ucla.edu; johnjn@ucla.edu; bjones@cse.ucla.edu

H. Cai
Guangdong University of Foreign Studies, Guangzhou, People's Republic of China
e-mail: hwcai@ucla.edu

H.F. O'Neil et al. (eds.), *Teaching and Measuring Cognitive Readiness*,
DOI 10.1007/978-1-4614-7579-8_14,

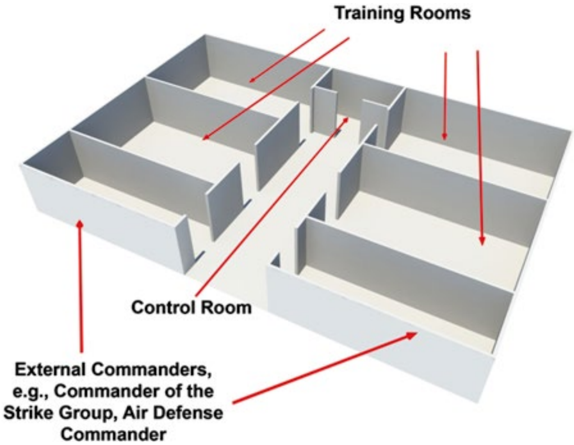


Fig. 14.1 Multi-Mission Tactical Trainer (MMTT) facility layout

tactical picture, making sound tactical decisions even with limited information, effectively communicating, prioritizing (threats, actions, and requirements), and thinking ahead. The focus is on individual training while also touching on the team concepts; the goal is for the SWOS DH graduates to join the training program on their ship and in a short period of time qualify for TAO.

One of the most important roles played by all Surface Warfare Officer Department Heads, in addition to their department duties, is to stand watch in the CIC as the TAO. The TAO is responsible for tactical employment and defense of the ship. He or she manages use of the ship's weapons and sensors, directs the movements of the ship, interprets rules of engagement, and monitors the movements and actions of friendly and enemy ships, planes, missiles, and submarines in the region. The TAO integrates this information to form a tactical picture of the situation, selects appropriate responses, issues orders, and informs the commanding officer of intentions and actions. Referring back to the Fletcher (2004) definition, this is clearly an example of performance in a complex and sometimes unpredictable environment requiring cognitive readiness.

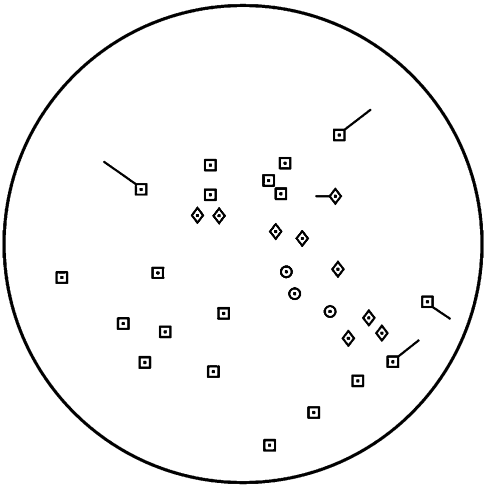
14.2 The Multi-Mission Tactical Trainer

An important part of SWOS's TAO training is practice and testing in a simulation facility called the Multi-Mission Tactical Trainer (MMTT) (U.S. Navy, 2012). The MMTT consists of four training rooms simulating CICs, two rooms for personnel playing the roles of commanders external to the ship, and one simulation problem control room arranged as shown in Fig. 14.1.

Students playing the role of the TAO and other CIC team members sit at personal computers in each training room arranged approximately as shown in Fig. 14.2.

Fig. 14.2 MMTT room layout

Fig. 14.3 Example MMTT display



Each student views a display like the example in Fig. 14.3. The icons represent tracks of sensed objects that may be friendly or hostile, or unknown airplanes, surface ships, or submarines. The vectors indicate direction and speed of movement.

The CIC team must identify each track, evaluate the threat posed by tracks declared hostile, and take action to defend against attack by hostile tracks. CIC team members are responsible for providing data to the TAO to support his or her situation assessment and decision making and performing actions as directed by the TAO. For example, the Anti-Air Warfare Coordinator monitors air tracks, informs the TAO of potential threats, and issues queries and warnings to unidentified tracks as ordered by the TAO. Table 14.1 summarizes the duties of all watchstanders in Fig. 14.2.

Table 14.1 Watchstander duties

Watchstander	Duties
Radar Systems Coordinator	Operates radar systems for detection and tracking of aircraft
Anti-Air Warfare Coordinator	Monitors aircraft, informs the TAO of potential threats, and issues queries and warnings to unidentified tracks
Tactical Information Coordinator	Operates the Tactical Digital Information Link, which communicates tactical data to friendly ships and aircraft
Missile Systems Supervisor	Responsible for use of surface-to-air missiles
Tactical Action Officer	Leads the CIC watch team and is responsible for defense of the ship
Combat Systems Coordinator	Responsible for the activation, monitoring, and deactivation of the combat systems that support the CIC
Electronic Warfare Supervisor	Operates the electronic emissions detection equipment used to detect and classify aircraft based on their radar signal emissions

TAO performance assessment was based on whether or not certain actions occurred (e.g., did the TAO order queries and warnings, did the TAO send an airplane to visually identify a suspected hostile track, or did the TAO shoot at a threatening hostile track?). This is considered the "what" of performance in the MMTT, the simulated CIC used at SWOS for training and assessment. In addition to these measures, SWOS was concerned with measuring the TAO's cognitive readiness, or the "why" of performance—the thinking behind the actions.

14.2.1 The MMTT Simulation Scenarios

Students receive a series of 21 training scenarios over 12 weeks, with each scenario requiring 30–45 min to complete. The training scenarios include problems involving air defense (AD), surface warfare (SUW), and undersea warfare (USW) threats, scenarios combining two or all three threat types (multi-threat), scenarios requiring coordination with other ships (coordinated operations), scenarios requiring operations in compressed (restricted area) battle spaces, scenarios focusing on communications, and a scenario presenting a problem focused on USW target motion analysis (requiring determination of the location, course, speed, and bearing of a submarine).

The 21 training scenarios are followed by two final test scenarios; if the student fails the first test scenario, he or she receives the second final test scenario. In the rare case that a student fails the second test scenario, he or she would receive remedial instruction and would be retested with a new scenario. The two final test scenarios are compressed battle space and multi-threat scenarios, with all three battle spaces in each: AD, USW, and SUW. The student must pass a test scenario to successfully complete the course.

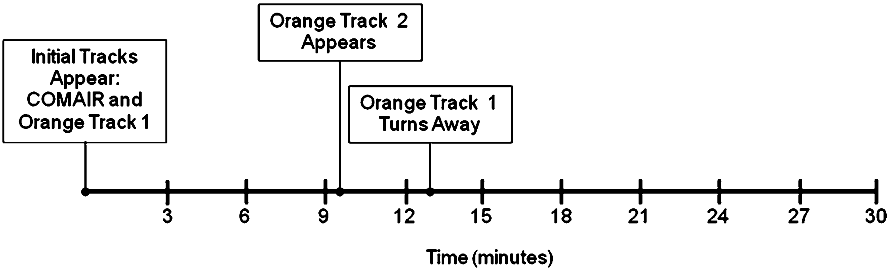


Fig. 14.4 Example MMTT simulation scenario timeline

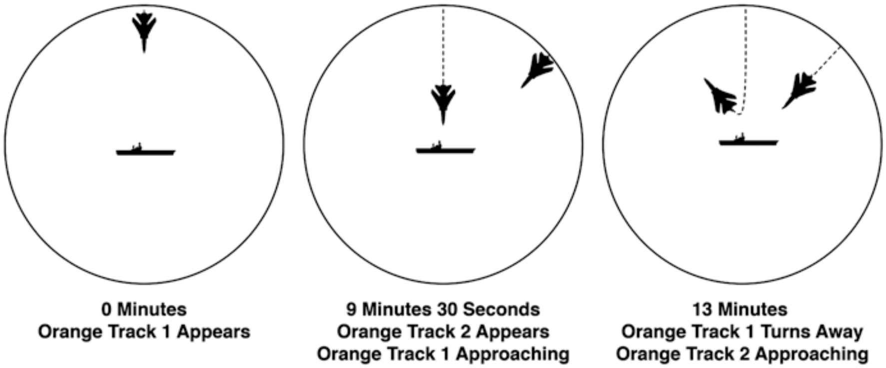


Fig. 14.5 Example MMTT simulation scenario event timing on the display

14.2.2 An Example Scenario

We cannot publish the MMTT scenarios actually used in training and testing as they are classified, but we can describe an example to provide a general idea of the contents of a scenario. Figure 14.4 summarizes a timeline for an air defense task scenario that might be presented in a MMTT session, and Fig. 14.5 shows the major events of this scenario roughly as they would appear on the MMTT displays as the scenario progresses.

The scenario opens with a display of several aircraft tracks, including commercial aircraft (COMAIR) and a track from a potentially hostile country called Orange (Orange Track 1). The CIC team begins to identify each track using information such as responses to automated electronic interrogation, commercial flight plans and times, formations, flight profile, direction, electronic emissions, altitude, distance, and speed. Orange Track 1 does not respond to electronic interrogation or radio queries and warnings, and it appears to be coming from country Orange, a potential enemy, so the TAO orders an airplane to intercept in order to obtain a visual identification. The visual identification will indicate whether the track is a country Orange airplane, and if so, whether it is a threat to the ship.

While Orange Track 1 is being investigated, Orange Track 2 appears, and the TAO must decide whether to continue investigating Orange Track 1 or reassign the intercepting airplane to Orange Track 2. Later in the scenario, Orange Track 1 is found not to be a threat as it turns away. The threat is Orange Track 2. Possible TAO errors are failure to issue timely queries and warnings to both tracks, failure to send the airplane to investigate Orange Tracks 1 and 2, and failure to initiate timely actions to defend against an attack by Orange Track 2.

14.2.3 Scenario Cognitive Demands

SWOS instructors designed the scenarios to elicit performance that would provide evidence of the examinee's knowledge, analytical ability, and decision-making skills. Each scenario is aligned with objectives, with scenario events serving as triggers to elicit performance providing evidence for mastery of the objectives. In addition to the events, the cognitive demand of each scenario is determined by the environment, including the pace of events, the political situation, and the geography, e.g., a compressed battle space with potential threats coming from several directions and requiring fast decisions is much more challenging than working in the open sea.

To obtain a measure of the relative cognitive demand of each scenario, we asked eight MMTT instructors, none of whom were involved in designing the scenarios, to rank order the difficulty of the 21 training scenarios and two final test scenarios. Criteria used to judge difficulty included number of potential threats, number of warfare areas, pace and complexity of the scenario, existence of standard operating procedures based on preplanned responses (PPRs), political tensions, and rules of engagement. To convert the rankings to an interval scale, we fit a Partial Credit Model (PCM), a Rasch model for polytomous data, to the rankings of the eight raters (see Harwell & Gatti, 2001, for the procedure to rescale ordinal data using item response theory models).

Prior to fitting the PCM to the data, we tested the unidimensionality and local independence assumptions (Embretson & Reise, 2000) through maximum-likelihood factor analysis in PASW 17.0. For the single factor model, $\chi^2 = 31.39$, df = 20, $p > 0.05$, there was evidence for the unidimensionality of the eight rankings. The result also excluded the concern over local dependence, as this single factor had an eigenvalue of 7.21, accounting for 90.1 % of the total variance. With these assumptions met, the PCM was fitted through the "gpcm" command in the "ltm" package of R2.9.1 (R Development Core Team, 2009).

The PCM produced standardized complexity values in the range between −3.19 and 2.05, with a mean of zero. To make them reader-friendly, these values were adjusted by adding a constant (4.19) so that the difficulty of the easiest scenario is 1.00. As a result, the most difficult scenario had a scale value of 6.24. In general, the communication scenarios, which were more concerned with using correct message format and content than tactics, were rated lowest in difficulty. The single threat scenarios (air defense, SUW, and anti-submarine warfare (ASW), here equivalent to

USW, or undersea warfare) were next, with ASW scenarios slightly more difficult than AD and SUW scenarios. The coordinated operations and multi-threat scenarios were more difficult than the single threat scenarios, and similar to one another in difficulty. Most difficult were the compressed battle space scenarios. The two final scenarios were rated as similar in difficulty to the multi-threat scenarios (Final 1) and compressed battle space scenarios (Final 2). Although Final 1 had a scaled difficulty lower than Final 2, i.e., 3.58 vs. 3.94, the difference was not significant, suggesting that the two test scenarios are indeed equivalent.

14.3 Measuring Cognitive Readiness

Revisiting Fletcher (2004), we propose that cognitive readiness is needed to perform competently in a complex and unpredictable environment, and we have suggested TAO performance in the CIC is an example. So we know it when we see it, but how do we measure it? We need a more detailed definition, and we need to operationalize it. Like intelligence and motivation, cognitive readiness is a theoretical concept, a characteristic of human performance that is latent and not directly observable, a construct invented to explain some observable phenomenon such as TAO performance in the MMTT. Furthermore, the concept is composed of several parts. O'Neil, Lang, Perez, Escalante, and Fox in this volume, for example, have identified 12 components, and other chapters in this volume, e.g., the chapter by Fletcher and Wind, propose lists with more or less similar components.

14.3.1 The Constructs

Due to lack of general agreement on how to decompose cognitive readiness, we felt justified in developing our own set of constructs, based on a review of the literature and our assessment of the appropriateness of components in studying TAO performance in the MMTT and meeting the assessment needs of the SWOS staff. We based our construct list on the constructs identified by Morrison and Fletcher (2001), the extensive situation awareness (SA) and sense-making literature (e.g., Blandford, 2004; Durso & Sethumadhavan, 2008; Endsley, 1995a, 1995b, 1997, 2000, 2004; Gaba, Howard, & Small, 1995; Gorman, Cooke, & Winner, 2006; Graham & Matthews, 2000; Klein, 1999; Klein, Moon, & Hoffman, 2006; McCarthy, 2006; Wickens, 2008), the intelligence analysis and cognitive bias literature (e.g., Bruce, 2008; Bruce & Bennett, 2008; Davis, 2008; George, 2004; Groopman, 2007; Heuer, 1999; Kerbel, 2004), and prior work on tactics planning and decision making (Cannon-Bowers & Salas, 1998; Morrison, Marshall, Kelly, & Moore, 1997; Radtke, Johnston, Biddle, & Carolan, 2007). We also paid close attention to the results of interviews with and feedback from the users of the assessment, the SWOS Department Head staff.

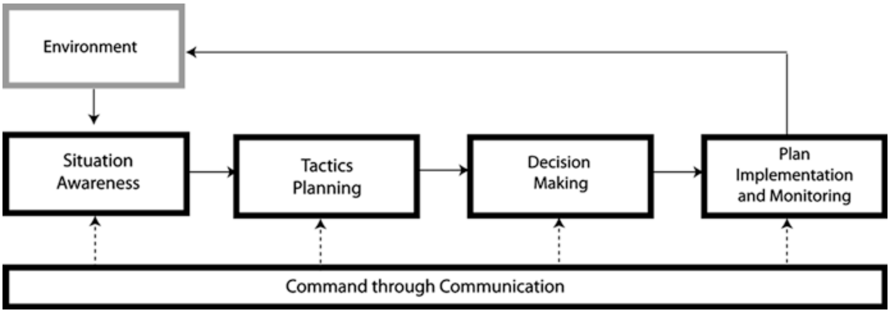


Fig. 14.6 A simplified overview of Tactical Action Officer (TAO) information processing

From our review of the literature and discussions with SWOS staff, we identified the following cognitive readiness constructs as being appropriate for studying TAO performance in the MMTT:

1. Situation awareness
2. Tactics planning
3. Decision making
4. Plan implementation and monitoring
5. Command through communication

The first four constructs are roughly in order of the information processing that takes place in the MMTT. The fifth applies to the entire information processing sequence. Figure 14.6 shows a simplified overview of how the process takes place. First, the TAO develops situation awareness to establish a rapid picture of the environment. This includes determining if a track poses a threat, which may involve predicting what the track is likely to do in the future. Once the situation is understood, the TAO needs to develop a tactical plan for responding to the situation. With a plan, the TAO must decide to execute the actions required to implement the plan, monitor its effects, and start the process again by assessing the impact of the effects on the situation. With respect to command and communication, the TAO is expected to act effectively and in a timely manner at every step of the process, and to communicate status, plans, intentions to act, and results of actions to the right people at the right time.

This description gives us the big picture of the cognitive readiness constructs. To develop measures, we need more detailed descriptions.

14.3.1.1 Situation Awareness

Of the five constructs, situation awareness (SA) is critical for successful performance in the MMTT. Most of the scenario time is spent on developing and maintaining SA, and in most scenarios the situation qualifies as a “complex and unpredictable environment of modern military operations," from the Fletcher (2004) definition. The other constructs in our list require planning and decision making, both good examples of what should be included in a list of cognitive readiness constructs. In the MMTT scenarios, however, because responding to a situation is often dictated by standard operating procedures, or PPRs, it is not emphasized. That is, given a situation, the TAO is expected to select and execute the appropriate plan, much as experts do in Klein's (1999) recognition-primed decision (RPD) model. This is not to say that these activities are not worthy of measurement, just that they are less important in the MMTT scenarios. For this reason, we will define SA in more detail than the other constructs.

Endsley (1988, 1997, 2000) defines situation awareness (SA) as "the perception of the elements in the environment within a volume of time and space, the comprehension of their meaning and the projection of their status in the near future" (1988, p. 97). The definition identifies three levels of situation awareness: perception, comprehension, and projection. Level 1, perception, refers to perception of "elements" in the environment, where elements are pieces of information that are meaningful in the domain, e.g., the altitude, speed, and heading of an aircraft. They are not mere stimuli, e.g., a vector on the screen. This is perception, not sensation.

Level 2, comprehension of the meaning of the perceived elements, refers to understanding the significance of the elements in light of operator goals (Endsley, 1995b, p. 37). For example, if the goal is to protect the mission essential unit (MEU), the altitude, speed, and heading of an incoming aircraft could be interpreted as a threat to the MEU (e.g., an aircraft carrier), which would lead to its classification as a contact of interest to be tracked closely, and possibly queried and warned.

Level 3, projection, is what might occur in the future to determine what course of action (COA) to take in the present. This is based on current status and hypotheses about a contact's intentions. If the political situation indicates possible hostile action and the contact is approaching at low altitude at high speed, for example, an attack would probably be predicted. If the contact circles near the limit of your weapons range, the approach might be interpreted as a test of our reaction, and no attack would be predicted.

Although SA may be viewed as a strictly linear process, with the perception of elements leading to comprehension of their meaning and the prediction of future action, it does not always occur in a stepwise fashion when SA operates in a complex, dynamic system such as the MMTT. The TAO's goals, comprehension (Level 2), and projections (Level 3) can focus the search for new information in Level 1.

As Klein's RPD model (Klein, 1999, 2008) suggests, experts match the current situation with situations experienced before and retrieve memory schema or mental models associated with the prior experience to accelerate comprehension, using the model to fill in elements that have not been perceived, creating expectations and identifying relevant cues to guide the search for new information, and identifying plans—goals and actions—that have been successful in the past in this situation. This speeds up comprehension, and makes situation awareness in a dynamic environment possible, but it can also cause errors due to forcing data into the model, as described in the cognitive bias literature, e.g., Groopman (2007) for medical

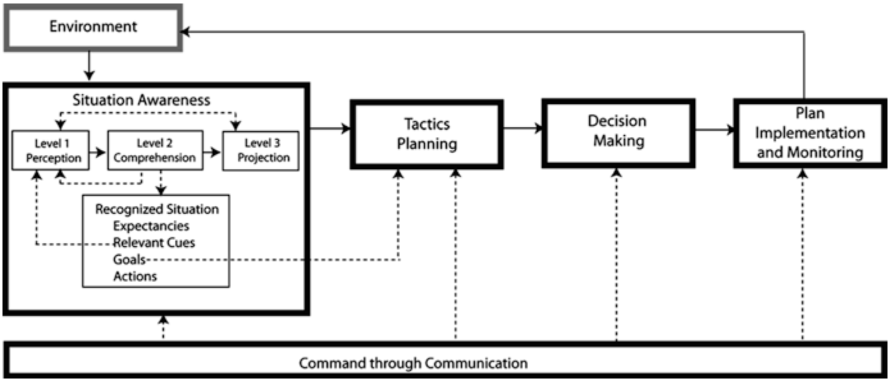


Fig. 14.7 An overview of TAO information processing with an expanded representation of situation awareness

decision making and several studies of cognitive errors in intelligence analysis, including CIA Directorate of Intelligence (1997), George (2004), and Heuer (1999). Because of this opportunity for error, Level 2 SA sometimes involves an effortful process of gathering, synthesizing, and evaluating evidence, described in the intelligence analysis literature as linchpin analysis (CIA Directorate of Intelligence, 1997). The sense-making literature is also relevant here (Leedom, 2001; Weick, 1995; Weick, Klein, Peluso, Smith, & Harris-Thompson, 2007), emphasizing the use of story-building to explain the observed evidence. Some see sense making as an alternative formulation of cognition in SA, but we agree with Endsley (2004) in determining that it is a subset of Level 2 processing.

Figure 14.7 expands upon the process that is illustrated in Fig. 14.6, by adding the three SA levels, and the interactions among functions that make TAO information processing a dynamic process using SA Level 2 products of the recognized situation (expectancies, relevant cues, goals, and actions) to influence perception (SA Level 1) and tactics planning. The three levels of situation awareness iterate as many times as is needed based on changes in the environment due to external causes or TAO actions.

14.3.1.2 Tactics Planning

The TAO must use situation awareness to determine potential COA. The goal in the CIC is tied to the primary mission of protecting the MEU, the ship most essential to successful completion of the mission (e.g., aircraft carrier). The tactics plan involves utilizing resources (e.g., data gathering through sensors and real assets such as helicopters and airplanes), and considering the reaction of the unknown track to friendly force action (e.g., responding or not to queries and warnings) in order to develop defensive and/or offensive action plans based on the rules of engagement.

As mentioned in the discussion of SA, when the situation is recognized as matching a situation experienced before, the TAO may not have to develop a plan (Klein, 1999). Rather, he or she can select a plan that has been taught in school or used successfully in prior situations, or if the situation is similar, but not identical, to a situation experienced before, the plan can be edited and tested by running a mental simulation.

14.3.1.3 Decision Making

Decision making is the process of choosing from alternative plans. As noted above, if the situation is recognized, the successful plan connected with that situation will be selected, or a plan associated with a similar situation will be edited and tested in a mental simulation (Klein, 1999, 2008). If the situation is not recognized, however, plans will be developed and alternative plans will have to be evaluated and one of them chosen. This may involve using some mental simulations or "what if" thinking.

14.3.1.4 Plan Implementation and Monitoring

Effective implementation involves making sure that the chosen COA is carried out as planned, and to maintain or revise the plan based on new information from the results of the implementation. If the situation takes a turn for the worse, then a contingency, or modified or alternative plan, must be put into place. In the MMTT, the orders given by the TAO and queries to the appropriate watchstanders (other members of the CIC team) along with the monitoring of the situation are relevant indicators of this construct.

14.3.1.5 Command Through Communication

This construct deals with commanding through proper communication of the TAO with others, including other watchstanders and superiors as well as other team members communicating with the TAO. It includes having the correct structure (following protocol and having the right format), treatment (message content, accuracy, delivery style/presentation), and timeliness. Researchers (Achille, Schulze, & Schmidt-Nielsen, 1995) describe four aspects of effective communication, particularly in military environments:

1. Accuracy: unambiguous and proper use of terms
2. Terseness (brevity): especially important in heavy-use communication nets
3. Selective (relevance): only pertinent information
4. Identification (to guide attention): communicator identifies himself/herself by name and/or by role

The importance of effective communication cannot be overemphasized, given issues related to message interference due to overload of the communications network, messages “stepping on” each other when two people try to use the communication channel at the same time, and the difficulty of handling multiple and different communications coming into different sides of the TAO’s and others’ headsets, often simultaneously.

Additionally, following communication protocols enables the effective command and control required for performance of all CIC functions. Situation awareness, tactics planning, decision making, and tactics plan implementation also depend on effective communication.

14.3.2 Operationalizing the Constructs

We have described constructs to clarify the meaning of cognitive readiness in the MMTT, but they are still vague terms. To translate a vague term into something that can be measured, we need an operational definition, a specification of the process used to determine its presence and quantity.

As suggested by Tenney and Pew (2006), Jones and Endsley (2004), and many others, there are several ways to assess the cognitive processing behind actions:

1. Mid-scenario probe questions. One can pose questions during the scenario without pausing the activity. The advantage is a low memory requirement because the report is immediate, not delayed. The disadvantages are that the questions may interfere with performance due to interruption, and they may direct attention to things that might be otherwise overlooked.
2. Part-task anticipation. Questions are posed following short scenarios or scenario fragments presenting a situation. The advantages are that the assessment is brief and focused and the memory requirement is low. The disadvantage is that short trials are not representative of real situations and workloads.
3. Critical events. An event presenting a problem or anomalous situation is inserted in the scenario, and the assessor observes the response. This is the approach taken by Smith-Jentsch, Johnston, and Payne (1998) in an earlier study of TAO tactical decision making, and more recently by Radtke et al. (2007) in a study of pilot decision making in air combat. The advantages are a low memory requirement and no mid-scenario questions to intrude on the student’s task performance. Although this method elicits behaviors, or the “what” of performance, it does not provide information on the “why,” or the thinking behind the actions.
4. After-action review (AAR). This is the usual approach to assessment in military team situations. The assessor and team review and discuss performance at the completion of the simulation scenario. The advantage is that it’s not intrusive. The disadvantage is a very high memory requirement for students and assessors, although performance data collected during the scenario can be used to support the AAR, e.g., as done in the critical event study by Radtke et al. (2007).

Our approach is a blend of all four, designed to eliminate or minimize the disadvantages while retaining the advantages of each approach separately. We evaluated the complete scenario as a string of connected part-tasks, with probe questions at natural pauses between parts, and we included critical events in each part-task designed to elicit not only actions but also actions that expose the "cognitive errors" mentioned earlier. In addition to providing information on cognitive readiness, the approach reduces the interruption due to questions and the memory requirement by asking questions only during the pauses and by timing pauses so that there is little delay between the response and the question. It also avoids the risk of the question cueing different behavior because the questions are asked after the event and responses have occurred. And because the scenario fragments are strung together to form a long scenario, it avoids the problem of testing with unrepresentative situations and workloads.

14.3.2.1 The Assessment Tool

Our approach was to provide a job aid to the instructor/evaluator who also ran the AAR. The approach was implemented in an assessment tool, a computer-based program providing a series of screens mapped to the events of the MMTT scenario. It was used to assess the performance on the two MMTT finals as criterion-referenced summative tests.

Figure 14.8 shows a screen for an air defense event. Each screen contains items (statements or questions) linked to several descriptions of student actions or responses. In the figure, the items are: (1) What is this track? (2) How do you know? and (3) What are your expectations regarding this track? The descriptions of student actions or responses associated with each item are shown as rectangles grouped into three categories: (1) Optimal, the rectangles at the left; (2) Adequate, the rectangles in the middle; and (3) Other, the rectangles at the right.

Descriptions and groupings into categories were initially defined by one rater and then edited for accuracy and completeness with assistance from six expert raters. The rater recorded the student's responses and actions for each item by selecting the appropriate rectangles. When the rater clicked the "Submit" button at the lower right corner of the screen, the items were scored (see Sect. 14.4.4) and the next screen was displayed. Rectangles at the top of the screen (AD, SUW, ASW, All, and Score) were used to filter items, for example, show only air defense (AD), SUW, or ASW items, show all items, or view the score report of student performance. The tabs below these rectangles showed scenario events in chronological order, left to right, e.g., Pre-planning, COMAIR (Commercial Air) ID, Potential Threat Track 1, Potential Threat Track 2, and AAR (After Action Review).

The rater moved from event to event by selecting the tabs. Items that require the evaluator to prompt the student by asking a question verbally included "instructor prompt" written in parentheses. The tool produced a report summarizing results at the level of events and an overall score, and there was an authoring system that allowed instructors to develop new assessments. See Fig. 14.9 for an example of a report, and Fig. 14.10 for an example authoring system screen.

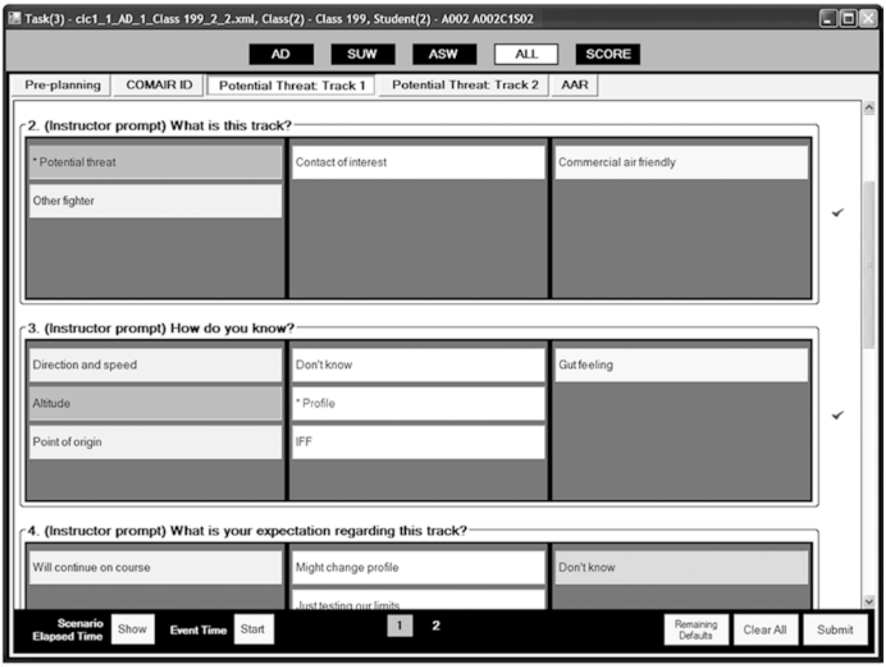


Fig. 14.8 An assessment tool screen

14.4 Data

14.4.1 Participants

We collected data from 427 Department Head students from eight classes. Most were ranked lieutenants (95.3 %), and most had 5–10 years of service in the Navy (56.6 %). A majority of the students were also most recently deployed to the Middle East (54.3 %), with the next highest percentage deployed to the Western Pacific (11.7 %). All but two had their deployment end within the last 7 years, and most (81.3 %) stated that their deployment supported ongoing operations. Approximately 88.6 % of the students in the first six classes had completed some Combat Systems Officer training, with the majority of this group (84.9 %) completing their training within the past 2 years. Two hundred eighty-eight students in these six classes were Combat Information Center Watch Officer qualified, with 83.7 % qualified in 2002 or 2003 (6 or 7 years before data collection). However, this information was unavailable for the other two classes due to a change in the background survey. A small percentage (13.1 %) of all students were TAO qualified.

Top title	Question	Label Text	Weight	subscore 1	subscore 2	Total
Pre-planning	(Instructor prompt) Based on the situation, what is the primary mi...	Building detailed Recognized Air Picture	2	2	2	4
	(Instructor prompt) What are you expecting to happen?	May need to utilize resources (F-18 to VID/escort/and engage as necessary)	2	2	2	
COMAIR ID	Checkprint and Evaluate	Checkprint and evaluate all COMAIR tracks in assigned reporting sector passing through VA	2	2	2	2
Potential Threat: Track 1	Track 1: Issue Queries	Query complete and Air reports to Bravo Whiskey while Track 1 is still at reasonable dist...	1	1	1	28
	(Instructor prompt) What is this track?	Potential threat	2	2	2	
	(Instructor prompt) How do you know?	Profile	1	1	1	
	(Instructor prompt) What is your expectation regarding this trac...					
	(Instructor prompt) What are your plans?	Send F-18 to VID	2	2	2	
	Send F-18	Escort until deemed non-threat or until directed to cease escort.	2	2	2	
	Evaluation of Track 1	Evaluate in CIEA prior to entry into VA, including review of VID/wings call.	1	1	1	
	Determine not hostile, orange aircraft	Never requested, hostile designation not required	2	2	2	
	Communication	Intentions are made clear internally only	1	1	1	
	Report	Report delivered at or shortly after entering the VA	1	1	1	
	Prompting	The report to Bravo Whiskey is made with prompting from the instructor	1	1	1	
	(Instructor prompt) What happened with the F-18?	F-18 ID as orange aircraft	2	2	2	
	(Instructor prompt) What does it mean?	orange aircraft only has air to air	2	2	2	
	(Instructor prompt) What is your expectation regarding this trac...	orange aircraft will continue on course	2	2	2	
	(Instructor prompt) What is your evaluation of the threat posed ...	A minor threat to other aircraft	1	1	1	
	(Instructor prompt) Why?	F-18 ID as orange aircraft	2	2	2	
	(Instructor prompt) What are your plans?	Call off F-18	2	2	3	
		Cover with own ships' missiles	1	1		
	Does the TAO effectively implement and/or oversee action plans?	TAO implements action plans adequately	1	2	2	
		TAO monitors action plan implementation adequately	1			
Potential Threat: Track 2	Track 2: Issue Queries	Query complete and Air reports to Bravo Whiskey while Track 1 is still at reasonable dist...	1	1	1	33
	(Instructor prompt) What is this track?	Potential threat	2	2	2	
	(Instructor prompt) How do you know?	Profile	1	1	1	
	(Instructor prompt) What is your expectation regarding this trac...					
	(Instructor prompt) What are your plans?	Send F-18 to VID	2	2	2	
	Reassign F-18	Send with enough time to intercept prior to entry into VA	1	2	2	
		Escort into and through VA or as directed by ADC.	1			
	Track 2: Issue Warnings	Issue warnings at adequate range	1	1	1	
	Evaluate the second hostile	Evaluate in CIEA prior to entry into VA, including review VID/wings call.	1	1	1	
	Communication	Intentions are made clear internally only	1	1	1	
	Report	Report delivered at or shortly after entering the VA	1	1	1	
	(Instructor prompt) What is your expectation regarding this trac...	Will continue on course	2	2	2	
	(Instructor prompt) What is your evaluation of the threat posed ...	A major threat	2	2	2	
	(Instructor prompt) Why?	F-18 ID as SU-27	2	2	2	
	(Instructor prompt) What are you	Continue escort, cover with F-18	1	1	1	

Fig. 14.9 An example report screen

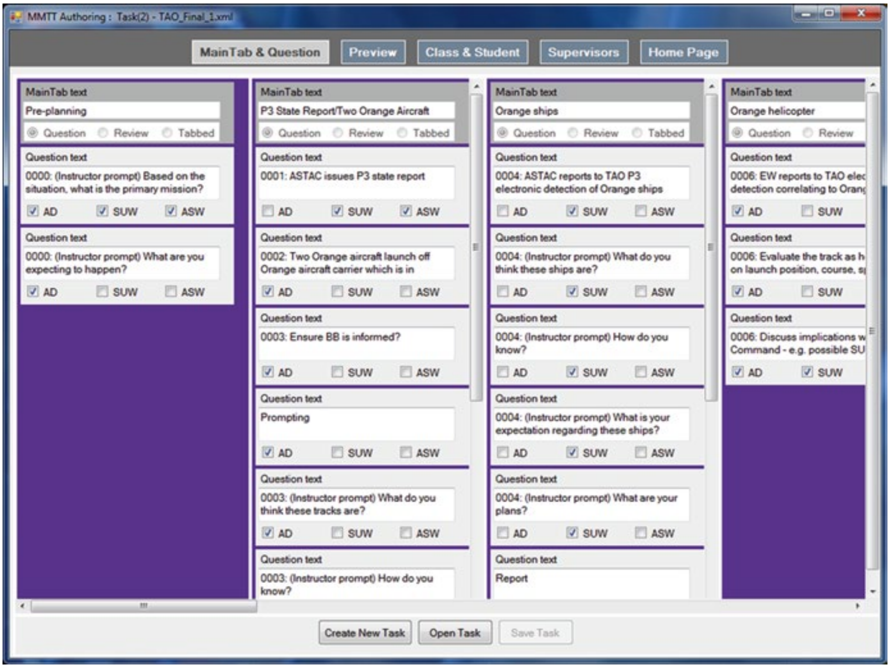


Fig. 14.10 An example authoring system screen

14.4.2 The Test Scenarios

Each student performed as a TAO on one of the two final scenarios. As noted above, both scenarios are multi-threat scenarios, with submarine, surface, and air tracks. Both scenarios take place in littoral environments, a compressed battle space, but Final 2 is more compressed than Final 1. The scenarios require between 30 and 45 min to complete.

To be able to present a student with a repeated test of the same knowledge and skill (e.g., presenting a second final test scenario when the student fails the first), or to be able to assess the performance of students receiving different final test scenarios against the same mastery standard, we need to be able to say that the scenarios are equivalent in terms of the knowledge and skills required for successful performance. As noted earlier, SWOS instructors rated all scenarios, and the results indicated that the two final scenarios were at the high end of the difficulty scale but not significantly different in difficulty.

Another test of equivalence is to compare examinee performance on the two test scenarios. We constructed an anchor test that contains a subset of the questions used in both test scenarios, on the basis of which the means and variances of the two finals were compared. Means and variances were not significantly different. With equal means and variances, the two forms are considered parallel measures (Lord & Novick, 1968).

14.4.3 Procedure

Department Head students received one of the two test scenarios for the TAO MMTT final examination. All voice messages were delivered through headphones, with different channels for each headphone, just as experienced by the TAO in a real CIC. One person played the role of all the watchstanders in the CIC, and triggered MMTT events defined by the scenario and based on the student's actions (e.g., issuing queries and warnings as directed by the subject). Subjects were given a pre-brief of the test situation the day before. During the test, a rater sat next to the student and recorded performance with the MMTT Assessment Tool using an adjacent desktop computer. A Post Commanding Officer from the SWOS staff was also in the room, playing the role of the Commanding Officer (CO), who also gave input to the scoring, and provided a more realistic simulation environment and increased the stress level to the student. The rater recorded all student actions and responses for each item. Each scenario included a preplanning tab, which asked the student what they perceived as being the important components of the mission (e.g., protect the MEU), and what they were expecting to happen. The scenario was then started. Students used their computer displays to view sensor data (including radars) and used a communications interface and headset for voice communications both internally and externally. At the end of the scenario, students were given a set of AAR questions and the rater used the assessment tool score report to provide the student with feedback on performance and the decision to pass or fail.

In addition to the total score, there was a tab in the assessment tool that kept track of single point failures, errors that a student made in the scenario that were costly in terms of lives, assets, or political relationships. For example, breaking rules of engagement, running a ship aground, or friendly fire would fit into this category. If a student received one or more of the "other" category in the single point failures tab, then the student failed the entire exam, regardless of their score on the rest of the items.

14.4.4 Scoring

Two points were awarded for each "Optimal" ranking and one point for "Adequate" responses, and −1 was given for the "Other" responses. Scores for all the responses for an item were summed to obtain the item score, and the item scores were used as the basis for further analyses. Construct scores were computed by adding the scores for the items mapped to each construct (e.g., situation awareness), and because possible scores were different for each construct, the scores were standardized by converting them to a percentage of the highest possible score as defined by expert raters.

14.4.5 Analysis

Two hundred twenty-seven students received the first test scenario (Final 1) and 200 received the second (Final 2). Students who did not pass were retested a day later with the alternative scenario. Twelve students failed the first time, eight on Final 1 and four on Final 2. There are two scores for these students, one for Final 1 and one for Final 2. All 12 students passed the second time.

We mapped the observed behaviors as recorded in the assessment tool to the constructs. For example, for situation awareness, we mapped the items that dealt with establishing the picture to Level 1 SA, questions such as “how do you know?” to Level 2 SA, and questions such as “what do you expect to happen?” to Level 3 SA.

The average for Final 1 was 140.19 out of a possible 178 (SD = 18.38) and for Final 2, 93.86 out of a possible 115 (SD = 12.04). The percentage of the mean to the maximum was close to 80 % for each final (78.8 % for Final 1 and 81.6 % for Final 2). The difference was not statistically significant ($z = 0.66$, $p = 0.49$). For Final 1 ($N = 227$) skewness is estimated at –0.474 with standard error 0.162, and kurtosis at 0.833 with standard error 0.322. Both statistics are within three standard errors. Kolmogorov–Smirnov statistic is estimated at 0.054 with DF = 227, $p = 0.20$. For Final 2 ($N = 200$) skewness is estimated at –0.474 with standard error 0.172, and kurtosis at –0.065 with standard error 0.342. Both statistics are within three standard errors. Kolmogorov–Smirnov statistic is estimated at 0.065 with DF = 200, $p = 0.042$. From these test statistics, we can say that the final scores are approximately normally distributed despite the high pass rates.

14.4.5.1 Technical Quality

The technical quality of the measures was determined by examining their validity and reliability. Construct validity refers to the degree to which assessment results can be interpreted as a meaningful measure of the construct for the purposes and situations to which they are applied. Although we intended to measure all five constructs identified in Sect. 14.3.1 (situation awareness, tactics planning, decision making, plan implementation and monitoring, and command through communication), we eliminated tactics planning and decision making from our analysis because, as noted above, responding to a situation in the MMTT is usually dictated by PPRs, so that the TAO is expected to select and execute the appropriate plan, not determine potential COA (tactics planning), and choose from alternative plans (decision making). And because most of what happens in a MMTT scenario is related to developing and maintaining Level 2 and Level 3 situation awareness, we split the situation awareness construct in two, one construct for Level 2 and another for Level 3. So we analyzed the data in terms of four constructs: CC = command through communication; SA2 = Level 2 situation awareness; SA3 = Level 3 situation awareness; and PI = plan implementation.

The two finals were developed to measure the same achievement and, as such, assume the same underlying constructs. However, the tests differ in four ways: the events involved, the items designed for each event, the number of items based on each event, and the number of items for the whole test. For the purpose of test validation, therefore, it is reasonable to construct a short form of each final by selecting items with essentially the same content in both finals as indicators of the constructs being measured. The following are examples of such items:

- How do you know?
- What are your plans?
- Communication
- Does the TAO effectively implement and/or oversee action plans?

These items were repeated in different events in both finals, and are regarded respectively as indicators of the SA2, SA3, CC, and PI constructs in both tests. Comprising such items, the short forms of the two finals are considered parallel forms of the same measure, and the structure of the parallel forms can be compared.

Because of practical considerations, we restricted the number of items in each short form to approximately 20. There were two such considerations. First, many items are worded differently in the two finals, and it is difficult to find more indicators of each construct with similar wording across the two finals. Second, due to the small sample sizes, it is difficult to test a statistical model with many variables. The short form of Final 1 included 21 items, with 5 items each for the SA2, SA3, and PI constructs, and 6 indicators for the CC construct. The short form of Final 2 included 19 items, with 4 items each for the SA2 and SA3 constructs, 6 indicators for the CC construct, and 5 indicators for the PI construct.

As the largest score range is eight among all the items used in both finals, we considered the scores for all the items to be ordinal variables. As Pearson correlation coefficients are underestimates of the relationship between such variables, the polychoric correlation matrix was used in subsequent analyses. The structure of both short form tests was analyzed through confirmatory factor analysis (CFA) in EQS6.1 (Bentler, 2007) using the reweighted least squares (RLS) estimator. After minor adjustments and replacements to improve fit, the four-correlated factor model was confirmed for both short form tests. For Final 1, the Satorra–Bentler scaled chi-square value (Satorra & Bentler, 1994) for the overall model fit was significant, $\chi^2(183)=274.40$, $p<0.05$, suggesting a lack of fit between the hypothesized model and the data. However, due to the sensitivity of χ^2, we based our judgment on other fit indices. Examination of these indices showed acceptable model fit with CFI = 0.967, RMSEA = 0.047, with the 90 % confidence interval between 0.035 and 0.058. As Table 14.2 shows, all four factors are significantly correlated with each other. Except for the relatively low correlation between SA3 and CC ($r=0.239$), all other correlations were moderate, in the range of 0.454 and 0.554.

All items loaded significantly onto their respective factors, with loadings ranging from 0.529 to 0.848. Table 14.3 shows the structure of the 21-item short form test.

Table 14.2 Correlation between factors for short form of Final 1

	SA2	SA3	CC
SA3	0.454*		
CC	0.536*	0.239*	
PI	0.466*	0.543*	0.554*

Note: *CC* command through communication, *PI* plan implementation, *SA2* Level 2 situation awareness, *SA3* Level 3 situation awareness
*$p<0.05$

Table 14.3 Factor structure for short form of Final 1

Items	CC	PI	SA2	SA3
Q34	0.765			
Q44	0.732			
Q33	0.730			
Q32	0.711			
Q46	0.699			
Q45	0.675			
Q10		0.795		
Q58		0.795		
Q19		0.777		
Q51		0.740		
Q38		0.709		
Q12			0.848	
Q28			0.833	
Q39			0.812	
Q27			0.744	
Q18			0.633	
Q07				0.791
Q54				0.740
Q41				0.615
Q08				0.562
Q29				0.529

Note: *CC* command through communication, *PI* plan implementation, *SA2* Level 2 situation awareness, *SA3* Level 3 situation awareness

For Final 2, the Satorra–Bentler scaled chi-square value (Satorra & Bentler, 1994) also indicated a lack of fit between the hypothesized model and the data, χ^2 (146) = 201.33, $p<0.05$, whereas the other fit indices indicated acceptable model fit, CFI = 0.934, RMSEA = 0.044, with the 90 % confidence interval between 0.027 and 0.057. The pattern of correlations between factors was similar to Final 1. As Table 14.4 shows, except for the correlation between SA3 and CC ($r=0.169$), all correlations between factors were significant, in the range of 0.317 and 0.730.

All items in this short form loaded significantly onto their respective factors, with loadings ranging from 0.289 to 0.861. Table 14.5 shows the structure of the 19-item short form test.

Table 14.4 Correlation between factors for short form of Final 2

	SA2	SA3	CC
SA3	0.607*		
CC	0.452*	0.169	
PI	0.730*	0.317*	0.408*

Note: *CC* command through communication, *PI* plan implementation, *SA2* Level 2 situation awareness, *SA3* Level 3 situation awareness
$*p<0.05$

Table 14.5 Factor structure for short form of Final 2

	CC	PI	SA2	SA3
Q17	0.796			
Q09	0.695			
Q10	0.633			
Q18	0.582			
Q35	0.514			
Q23	0.336			
Q12		0.789		
Q19		0.809		
Q24		0.861		
Q29		0.700		
Q37		0.634		
Q03			0.605	
Q11			0.574	
Q01			0.369	
Q13			0.315	
Q32				0.617
Q14				0.598
Q06				0.571
Q08				0.289

Note: *CC* command through communication, *PI* plan implementation, *SA2* Level 2 situation awareness, *SA3* Level 3 situation awareness

In addition to validity, we also looked at the reliability of the scores by construct. Reliability refers to the consistency of assessment results. Low reliability indicates that the scores are influenced by random errors such as the level of students' motivation or the inconsistency in rater scoring across students. Reliability also constitutes a necessary condition for construct validity. Cronbach's alpha (Pedhazur & Schmelkin, 1991) was obtained as an estimate of the reliability for each short form. Table 14.6 shows the alpha reliabilities by construct. For the short form of Final 1, most of them are relatively high, ranging between 0.707 and 0.811. For the short form of Final 2, the alpha values are lower, ranging from 0.310 to 0.808. The construct measured with the lowest reliability in the short form of Final 2 was SA2. As only four items are used in the analysis, we believe increasing the number of similar items may solve the problem of low reliability.

Table 14.6 Cronbach's alpha by construct for both short forms

	CC	PI	SA2	SA3
Final 1	0.727	0.811	0.707	0.723
Final 2	0.562	0.808	0.310	0.383

Note: *CC* command through communication, *PI* plan implementation, *SA2* Level 2 situation awareness, *SA3* Level 3 situation awareness

14.4.5.2 Preliminary Conclusions

What do these results mean? The CFA results provide preliminary evidence that both final test scenarios are valid measures of the four constructs we have identified, and the fact that the pattern of correlations among factors was similar for both scenarios in their short forms provides further evidence of similar factor structures.

The evidence for reliability is less comforting, especially for Final Scenario 2. We suspect that the problem is that there are an insufficient number of items mapping to each construct in Final Scenario 2. This scenario should be revised to increase the number of items for each construct, and to increase the similarity in wording of items for each construct. In addition to concerns about the scenarios, these results were based on a barely sufficient number of subjects according to the subject–variable ratio of 20:1 for factor analysis recommended by Hair, Anderson, Tatham, and Black (1995). Suppose we need to double the number of variables to increase the subscale reliabilities in further studies, then we would need data from at least 900 subjects. However, given the throughput in this course (maximum $n=480$), this approach would not be feasible.

14.5 Next Steps

This chapter described our approach and progress toward developing and validating an assessment of cognitive readiness. It is clearly a work in progress, with more to do in construct definition and assessment development and validation. But even with more to do, we think a description of our approach, results so far, the problems encountered, and our plans for next steps may be useful to others, and we offer it in that spirit.

As noted above, additional subjects are needed. The rate of data collection is slow because Department Head classes are usually composed of about 50–60 students and there are typically only five classes per year, but we continue to collect data. We also need better outcome measures than the Department Head Final Examination score, such as measures that are more directly rated to TAO performance in the CIC. SWOS has been conducting a questionnaire of Commanding Officers and Executive Officers, asking for a rating of performance of recent Department Head students as TAOs during their first cruise following SWOS.

The rate of data collection here is also slow, although it is ongoing. A better measure would be to use the MMTT Assessment Tool to assess performance of recent students as TAOs during their first cruise. We are investigating the possibility of completing this task on several ships at sea. Another option could be to gather data during a DH return visit to SWOS. Plans are to have Department Heads who have completed their first DH tour to return to SWOS for 2 days prior to moving on to their second tour. During this 2-day stop, the Officer will complete a Maritime Warfare assessment using MMTT, a shiphandling assessment, a 4-h written exam, and a leadership review. Each event is part of a process to obtain a Command Qualification (W. Chidester, personal communication, December 7, 2012).

The MMTT Assessment Tool would be improved by capturing TAO and other watchstander keystrokes so that the observer is relieved of the responsibility to record detailed user actions as well as responses to the cognitive readiness probes. The MMTT software records keystrokes, so the data exist. We are discussing approaches to capturing the keystroke data with SWOS and the MMTT software vendor. We believe that our probe items do not disrupt task performance, either by adding to workload or by changing behavior due to the expectation that probes will be presented. We plan to conduct studies testing these possibilities using methods based on those described by Jones and Endsley (2004).

Another possible area of research may be in analyzing the data using other methods. Traditional Bayesian network analysis (West et al., 2009), and a fully Bayesian approach using Markov Chain Monte Carlo, known as Personal Response Curves (Choi, Kang, & Delacruz, 2009) may provide alternative ways to analyze the data. The Bayesian networks may provide ways in which the issue of items mapping to multiple constructs may allow them to be incorporated rather than eliminated from analyses.

We also are concerned about rater reliability. We trained the initial raters in using the MMTT Assessment Tool and standards for judging performance. However, raters leave and are replaced by new raters who are trained by Department Head staff. We plan to assess rater standards and restart our training of TAO performance raters.

Finally, the major constraint in this research, and in all scenario-based training and assessment research, is the scenario design. The MMTT scenarios were designed to elicit performance providing evidence of the examinee's knowledge and skills, with scenario events triggering performance providing evidence for mastery of the objectives. However, the links between objectives and events, events and responses, and responses and competencies are not well defined. We cannot say that two scenarios are equivalent with certainty, or why they are different in a way that would allow us to predict or explain differences in performance. This is not peculiar to the MMTT scenarios; we believe it is a problem for all scenario-based training and assessment systems. We need a scenario-generation system that supports the design and development of complex training/assessment scenarios with task–competency and competency–evidence alignment and scenario branching based on embedded assessment measures. It is an understatement to say that this is a difficult task, but we consider it a requirement.

Acknowledgments The work reported herein was supported by grant number N00014-06-1-0711 from the U.S. Navy, Office of Naval Research with funding to the National Center for Research on Evaluation, Standards, and Student Testing (CRESST).

The findings and opinions expressed in this report are those of the authors and do not necessarily reflect the positions or policies of the U.S. Navy, Office of Naval Research.

References

Achille, L. B., Schulze, K. G., & Schmidt-Nielsen, A. (1995). An analysis of communication and the use of military terms in Navy team training. *Military Psychology: Special Issue: Team Processes, Training, and Performance, 7*, 95–107.

Bentler, P. M. (2007). *EQS 6.1 for windows [computer software]*. Encino, CA: Multivariate Software.

Blandford, A. (2004). Situation awareness in emergency medical dispatch. *International Journal of Human Computer Studies, 61*(4), 421–452.

Bruce, J. B. (2008). Making analysis more reliable: Why epistemology matters to intelligence. In R. Z. George & J. B. Bruce (Eds.), *Analyzing intelligence: Origins, obstacles, and innovations* (pp. 171–190). Washington, DC: Georgetown University Press.

Bruce, J. B., & Bennett, M. (2008). Foreign denial and deception: Analytical imperatives. In R. Z. George & J. B. Bruce (Eds.), *Analyzing intelligence: Origins, obstacles, and innovations* (pp. 122–137). Washington, DC: Georgetown University Press.

Cannon-Bowers, J. A., & Salas, E. (Eds.). (1998). *Making decisions under stress: Implications for individual and team training*. Washington, DC: American Psychological Association.

Choi, K. C., Kang, T., & Delacruz, G. C. (2009, July). *Intersection of psychometrics with games*. Presentation at the Psychometrics of Simulation Conference, Redondo Beach, CA.

CIA Directorate of Intelligence. (1997). *A compendium of analytic tradecraft notes: Volume 1, notes 3 and 4*. Retrieved from http://www.au.af.mil/au/awc/awcgate/cia/tradecraft_notes/contents.htm#contents

Davis, J. (2008). Why bad things happen to good analysts. In R. Z. George & J. B. Bruce (Eds.), *Analyzing intelligence: Origins, obstacles, and innovations* (pp. 157–170). Washington, DC: Georgetown University Press.

Durso, F. T., & Sethumadhavan, A. (2008). Situation awareness: Understanding dynamic environments. *Human Factors, 50*(3), 442–448.

Embretson, S. E., & Reise, S. (2000). *Item response theory for psychologists*. Mahwah, NJ: Erlbaum.

Endsley, M. R. (1988). Design and evaluation for situation awareness enhancement. In *Proceedings of the Human Factors Society 32nd Annual Meeting* (Vol. 1, pp. 97–101). Santa Monica, CA: Human Factors Society.

Endsley, M. R. (1995a). Measurement of situation awareness in dynamic systems. *Human Factors, 37*, 65–84.

Endsley, M. R. (1995b). Toward a theory of situation awareness in dynamic systems. *Human Factors, 37*, 32–64.

Endsley, M. R. (1997). The role of situation awareness in naturalistic decision making. In C. E. Zsambok & G. Klein (Eds.), *Naturalistic decision making* (pp. 269–283). Mahwah, NJ: Erlbaum.

Endsley, M. R. (2000). Theoretical underpinnings of situation awareness: A critical review. In M. R. Endsley & D. J. Garland (Eds.), *Situation awareness analysis and measurement* (pp. 3–32). Mahwah, NJ: Erlbaum.

Endsley, M. R. (2004). Situation awareness: Progress and directions. In S. Banbury & S. Tremblay (Eds.), *A cognitive approach to situation awareness: Theory and application* (pp. 317–341). Aldershot: Ashgate Publishing.

Fletcher, J. D. (2004). *Cognitive readiness: Preparing for the unexpected* (IDA Document D-3061). Alexandria, VA: Institute for Defense Analyses.

Gaba, D. M., Howard, S. K., & Small, S. D. (1995). Situation awareness in anesthesiology. *Human Factors, 37*, 20–31.

George, R. Z. (2004). Fixing the problem of analytical mind-sets: Alternative analysis. *International Journal of Intelligence and Counterintelligence, 17*, 385–404.

Gorman, J. C., Cooke, N. J., & Winner, J. L. (2006). Measuring team situation awareness in decentralized command and control environments. *Ergonomics, 49*, 1312–1325.

Graham, S. E., & Matthews, M. D. (2000). Modeling and measuring situation awareness. In J. H. Hiller & R. L. Wampler (Eds.), *Workshop on assessing and measuring training performance effectiveness* (Tech Report 1116) (pp. 14–24). Alexandria, VA: U.S. Army Research Institute for the Behavioral and Social Sciences.

Groopman, J. (2007). *How doctors think*. Boston: Houghton Mifflin.

Hair, J. F. J., Anderson, R. E., Tatham, R. L., & Black, W. C. (1995). *Multivariate data analysis* (4th ed.). Upper Saddle River, NJ: Prentice Hall.

Harwell, M. R., & Gatti, G. G. (2001). Rescaling ordinal data to interval data in educational research. *Review of Educational Research, 71*(1), 105–131.

Heuer, R. J. (1999). *Psychology of intelligence analysis*. Washington, DC: Center for the Study of Intelligence, Central Intelligence Agency.

Jones, D. G., & Endsley, M. R. (2004). Use of real-time probes for measuring situation awareness. *The International Journal of Aviation Psychology, 14*(4), 343–367.

Kerbel, J. (2004). Thinking straight: Cognitive bias in the US debate about China. *Studies in Intelligence, 48*(3), 385–404. Retrieved from http://www.au.af.mil/au/awc/awcgate/cia/think_cog_bias.htm

Klein, G. (1999). *Sources of power: How people make decisions*. Cambridge, MA: MIT Press.

Klein, G. (2008). Naturalistic decision making. *Human Factors, 50*(3), 456–460.

Klein, G., Moon, B., & Hoffman, R. R. (2006). Making sense of sensemaking 1: Alternative perspectives. *IEEE Intelligent Systems, 21*(4), 70–73.

Leedom, D. K. (2001). *Final report: Sensemaking symposium*. Vienna, VA: Command and Control Research Program, Office of the Assistant Secretary of Defense for Command, Control, Communications and Intelligence.

Lord, F. M., & Novick, M. R. (1968). *Statistical theories of mental test scores*. Reading, MA: Addison-Wesley.

McCarthy, G. (2006). Situation awareness in medicine. *Quality & Safety in Health Care, 15*, 384.

Morrison, J. E., & Fletcher, J. D. (2001). *Cognitive readiness* (IDA Paper P-3735). Alexandria, VA: Institute for Defense Analyses.

Morrison, J. G., Marshall, S. P., Kelly, R. T., & Moore, R. A. (1997). Eye tracking in tactical decision-making environments: Implications for decision support. In *Proceedings of the 1997 Command and Control Research and Technology Symposium* (pp. 355–364). Washington, DC: National Defense University.

Pedhazur, E. J., & Schmelkin, L. P. (1991). *Measurement, design, and analysis: An integrated approach* (studentth ed.). Hillsdale, NJ: Erlbaum.

R Development Core Team. (2009). *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing. Retrieved from http://www.R-project.org

Radtke, P., Johnston, J. H., Biddle, E., & Carolan, T. F. (2007, December). Integrating and presenting performance information in simulation-based air warfare scenarios. In *Proceedings of the Interservice/Industry Training, Simulation and Education Conference*, Orlando, FL.

Satorra, A., & Bentler, P. M. (1994). Corrections to test statistics and standard errors in covariance structure analysis. In A. von Eye & C. C. Clogg (Eds.), *Latent variables analysis: Applications for developmental research* (pp. 399–419). Thousand Oaks, CA: Sage.

Smith-Jentsch, K. A., Johnston, J. H., & Payne, S. C. (1998). Measuring team-related expertise in complex environments. In J. A. Cannon-Bowers & E. Salas (Eds.), *Making decisions under stress* (pp. 61–87). Washington, DC: American Psychological Association.

Tenney, Y. J., & Pew, R. W. (2006). Situation awareness catches on: What? So what? Now what? In R. C. Williges (Ed.), *Reviews of human factors and ergonomics* (Vol. 2, pp. 1–34). Santa Monica, CA: Human Factors and Ergonomics Society.

U.S. Navy, Surface Warfare Officers School Command. (2012). *Multi-Mission Tactical Trainer (MMTT), Tactical Action Officer (TAO) Intelligent Tutoring System (ITS).* Retrieved from https://www.netc.navy.mil/centers/swos/Simulators.aspx?ID=2

Weick, K. E. (1995). *Sensemaking in organizations*. Thousand Oaks, CA: Sage.

Weick, W. R., Klein, G., Peluso, D. A., Smith, J. L., & Harris-Thompson, D. (2007). *FOCUS: A model of sensemaking* (Technical Report 1200). Arlington, VA: United States Army Research Institute for the Behavioral and Social Sciences.

West, P., Rutstein, D. W., Mislevy, R. J., Liu, J., Levy, R., DiCerbo, K. E., et al. (2009, June). *A Bayes net approach to modeling learning progressions and task performances.* Paper presented at the Learning Progressions in Science (LeaPS) Conference, Iowa City, IA.

Wickens, C. (2008). Situation awareness: Review of Mica Endsley's 1995 articles on situation awareness theory and measurement. *Human Factors, 50*(3), 397–403.

Chapter 15
Software Support for Teaching and Measuring Cognitive Readiness

Allen Munro

15.1 Cognitive Readiness for Tasks

People are more or less well prepared to perform particular tasks. Someone who has never heard of the game of chess, who has never seen a game, and who has not been told how to play the game is clearly not cognitively ready to play chess. The person might be cognitively ready to *learn* to play chess, but that is a different matter. On the other hand, someone who has played in many tournaments and who has attained Master status is cognitively ready to play chess at a very high level, at least so long as that person is not cognitively impaired by drugs, alcohol, or some other agency.

Does it make sense to speak about cognitive readiness outside of the context of a task? That is, can we say that person A is in some sense "cognitively ready" and that person B is not cognitively ready, no matter what the task? Someone who suffers from severe intellectual deficits is likely not to be cognitively ready for most tasks. But this is not a very useful way to use the term cognitive readiness, because there are already other constructs, such as *intelligence*, that seem to fill the bill well enough. Someone might be an expert chess player or an outstanding military tactician, but that does not mean that they are cognitively ready to compose a symphony.

A. Munro (✉)
Center for Cognitive Technology, Rossier School of Education, University of Southern California, 250 North Harbor Drive, Suite 309, Redondo Beach, CA 90277, USA
e-mail: munro@usc.edu

H.F. O'Neil et al. (eds.), *Teaching and Measuring Cognitive Readiness*,
DOI 10.1007/978-1-4614-7579-8_15, © Springer Science+Business Media New York 2014

Cognitive readiness must be restricted to types of tasks. (O'Neil, Lang, Perez, Escalante, and Fox in this volume describe an approach in which some cognitive readiness attributes are treated as being relatively task independent.)

15.2 Levels of Cognitive Readiness

When it is said that someone is cognitively ready to perform some type of task, there is an ambiguity about the level of skill or expertise that they will exhibit when performing examples of that task. One person might be cognitively ready to perform a simple example of the task, but not be able to perform well on a more complex instance of that type of task. Such a person could be said to have a moderate level of cognitive readiness for that task. Another person might have expert knowledge about the task and all its variations. When motivated to engage in the task, that person would exhibit a high degree of cognitive readiness for the task.

15.2.1 Levels of Cognitive Readiness in Complex Decision-Making Tasks

For several years, I have been involved in research on learning decision-making and problem-solving skills in the US Navy contexts. The structure of these real-world tasks leads me to propose five levels of cognitive readiness for such tasks. In the following list, the lowest level or simplest type of cognitive readiness appears first on the list, and the most complex type appears last.

- Cognitive readiness for categorization
- Cognitive readiness to act on the basis of a categorization
- Cognitive readiness to accommodate special conditions in procedure execution
- Cognitive readiness for task complexity (such as the occurrence of instances of multiple concepts, which requires prioritizing or combining procedures)
- Cognitive readiness to generate new types of solutions to novel problems

The "categorization" level of cognitive readiness can be seen as related to situational awareness (Endsley, 1995). In situations in which a decision-maker perceives relevant events and comprehends their import well enough to label the situation correctly, the decision-maker has used situational awareness to usefully categorize the situation.

Someone who knows what to do and how to do it, based on the categorization, can be said to be cognitively ready to act. The two higher levels of cognitive readiness listed above require the application of metacognitive strategies.

15.2.1.1 The *TAO Sandbox*[1]: A Problem-Solving Task

Tactical Action Officers (TAOs) are Navy surface fleet Department Heads who must coordinate information and recommend tactics to the captain during tactical engagements. The TAO Sandbox is based, in part, on an earlier research product that was designed to support TAO planning for antisubmarine warfare (ASW), called the ASW Sandbox (Auslander, Molineaux, Aha, Munro, & Pizzini, 2009; Munro & Pizzini, 2009; Munro, Pizzini, & Bewley, 2009). The TAO Sandbox extends that earlier product in many ways, but principally by supporting TAO planning for Air Defense and for Surface Warfare. In both the TAO Sandbox and its predecessor, a problem-solver can develop a plan and then let simulated time pass to see how the plan plays out. Plans can be reformulated and changes made during the course of a session. As the student uses the Sandbox, every action is recorded in a file. Sessions can be replayed in the Instructor Mode of the Sandbox, which is also used to develop new scenarios.

The TAO Sandbox has been adopted for a variety of uses in the training of TAOs, in the Department Head course at the Surface Warfare Officers School (SWOS). Instructors and students use the TAO Sandbox in a number of ways, including these:

- Instructors build simple scenarios to illustrate particular tactical concepts.
- Instructors illustrate these concepts by using the scenarios in class. In some cases, they record a scenario session in advance and then use the playback feature to present the recorded session, while providing commentary.
- Students solve illustrative tactics problems, often in small groups.
- Students present their tactical solutions, debriefing in the context of a recording of their sessions.
- Some highly motivated TAO students experiment with the Sandbox, building their own scenarios in different combat contexts, such as ASW, surface warfare, or air defense.
- The Sandbox is sometimes used to help assess a student's knowledge of tactics.

SWOS management has identified potential areas for additional utilization of the TAO Sandbox, including the training of littoral operations, which are close to shore, and expeditionary warfare tactics and procedures, which involve landing and supporting ground forces from the sea.

[1] The term Sandbox is an analogy to the physical sandboxes, usually raised on legs, that are used to teach tactical principles to Army and Marine officers, where objects representing tanks, firing teams, squads, platoons, artillery pieces, etc. are placed in terrains modeled using the sand in the box. Our computer-based Sandbox for teaching about Navy tactics allows placements of naval units to model tactical situations. In addition, however, it can simulate the passage of time and automate the movement of units.

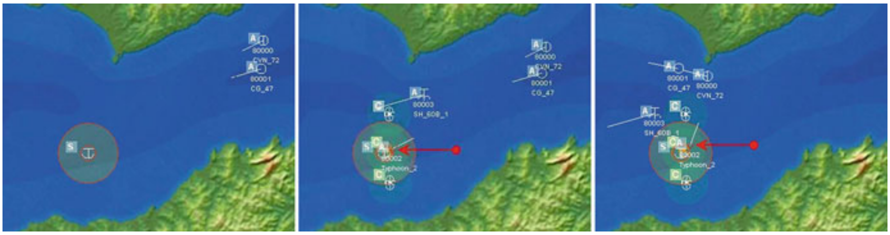


Fig. 15.1 Three states of an imagined transit scenario

The above-listed uses of the sandbox have replaced a number of less interactive methods for presenting and practicing tactics, including the use of static sketches on a white board and the use of canned animations or videos. When assessments are conducted in the context of the TAO Sandbox, the student produces a set of actions that have effects in the simulated tactical areas, rather than writing a short essay about a selected tactical approach and its expected outcomes.

TAO Sandbox "problems" are scenarios that students can "solve" in a variety of ways that can include deploying various sensors to detect potential hostile units and positioning one's own ships to avoid threats, or in such a way that more vulnerable units are protected by more capable units, by issuing appropriate queries and warnings to potentially hostile units, by targeting and attacking hostile units when appropriate, and so on. The combination of actions that constitute an acceptable solution [as judged by human experts in after-action reviews (AARs), for example] will vary, depending on the initial conditions of the problem and on the autonomous actions taken by hostile and neutral units during the course of a session.

One of the advantages of the Sandbox over previously utilized methods is that it does not rely on *mental* simulation of expected results. Working memory and other processing demands make it difficult to accurately predict the changes in relative positions of the involved ships, submarines, and aircraft over time. The built-in logic of the Sandbox handles these effects, and others, such as automated hostile behaviors on the part of simulated opponents, as conditions warrant during a session.

Cognitive load theory is, according to Paas, Renkl, and Sweller (2004), "concerned with the learning of complex cognitive tasks, where learners are often overwhelmed by the number of information elements and their interactions that need to be processed simultaneously before meaningful learning can commence" (p. 1). The theory assumes a working memory of limited capacity. When one attempts to solve a complex tactics problem by mentally simulating the interactions of a number of ships over time, working memory limitations may lead one to oversimplify or to ignore potential interactions between even simple events, such as the independent movement of different ships and how that will affect whether those ships will enter the ranges of each other's weapons systems.

Figure 15.1 represents possible states of a mental simulation of a transit that is threatened by a reported hostile submarine. The circle at the lower left is the Torpedo Danger Area (TDA), an expanding area of risk that may contain the hostile submarine.

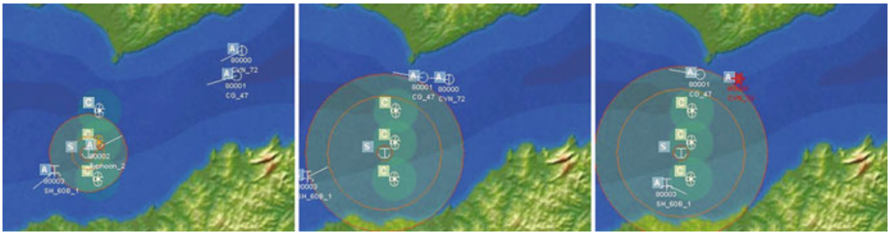


Fig. 15.2 Three states of a simulated scenario

The first state illustrates the initial condition. The two friendly ships (a cruiser labeled CG_47 and an aircraft carrier labeled CVN_72), which are on the right, are to proceed to the west through the Strait of Gibraltar. There is a concern with the report of a possibly hostile submarine near the center of the TDA.

In the second state shown in the figure, the tactician imagines having launched a helicopter that carries sonobuoys and having placed three of the sonobuoys in a line crossing the TDA. He imagines that one or more of the sonobuoys will detect the hostile sub. In this part of the figure, an arrow points to the detected submarine position.

In the third state shown in Fig. 15.1, the tactician imagines that some time later, the sonobuoys are still tracking the hostile submarine and the friendly ships are successfully avoiding it and completing their transit of the strait.

When the same exercise is carried out in the context of the TAO Sandbox, however, the previously imagined results of the plan are shown to be incomplete and inaccurate.

The first part of Fig. 15.2 shows the results of a tactician actually attempting the problem solution that was merely imagined in the illustrations of Fig. 15.1. Sonobuoys have been dropped, and the hostile sub has been detected.

In the second part of the figure, we see that before the friendly units have exited the strait, the submarine has exited the zone of detection of the deployed sonobuoys. Apparently the *mental* simulation shown in Fig. 15.1 did not accurately reflect the relative speeds of the submarine and the friendly units.

In the third part of Fig. 15.2, the submarine has successfully launched a torpedo and destroyed the aircraft carrier, which was undoubtedly the Mission Essential Unit (MEU). (The MEU is the unit, typically an aircraft carrier or large amphibious ship, which must be protected, because the assigned mission cannot be carried out without it.)

When the tactician using the TAO Sandbox attempts to carry out the imagined plan, it doesn't take long to find out that the submarine escapes detection long before the friendly ships reach the potential danger zone. The tactician is operating in the dark and may be placing his or her ships at risk. So the Sandbox user restarts the scenario and tries a different solution, as shown in Fig. 15.3.

This time, when the probable course of the submarine is determined, additional sonobuoys are dropped along that course, so that the sub is continuously tracked. (The tactician also chooses a sonobuoy type with a longer battery life.) See the first part of Fig. 15.3. Because the sub is moving toward the north, the friendly ships are

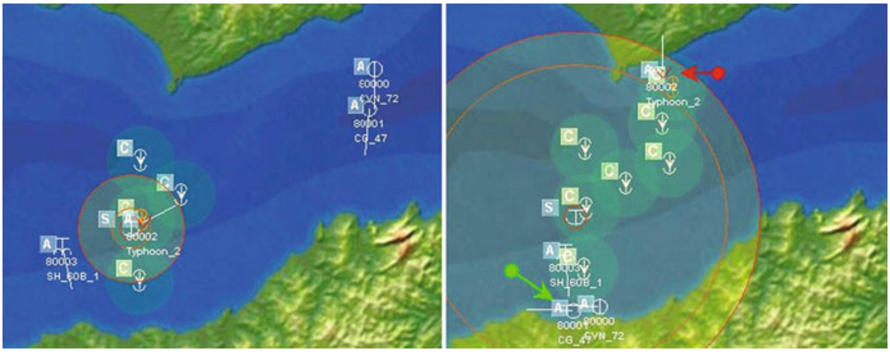


Fig. 15.3 Another attempt—same scenario

turned south and then back to the west, so that they will pass well to the south of the torpedo range of the sub in its known position. See the second part of Fig. 15.3. Here the arrow above marks the tracked location of the submarine; the arrow at the lower left marks the position of the two friendly ships that have successfully evaded the hostile sub.

In addition to providing a more accurate practice environment than mental simulation, the TAO Sandbox offers a number of other features that promote learning. There is considerable evidence that opportunities to practice the application of knowledge with guidance in a variety of environments promote learning and the acquisition of expertise. This work has been reviewed and summarized very effectively by Clark, Yates, Early, and Moulton (2010) extending earlier work by Merrill (2002). Mayer (2004) also cogently reviews related studies that show the importance of combining direct instruction with practice. Practicing certain tactics, such as ASW, in real-world environments on board ships, dueling with realistic opponents, is a very slow, very expensive form of practice. Given the speeds at which submarines can operate in relative silence, ASW tactics can take the form of slow motion duels. A TAO can reasonably be exposed to only a very few such real-world exercises, both because of the time they consume and because of the great expense of devoting entire crews and many hours of ship time to carrying out the selected tactics. The TAO Sandbox, on the other hand, has features that enable time to be greatly speeded up. Some ASW scenarios can be conducted at up to 850 times the rate that it would take to carry them out in the real world. This means that a prospective TAO can be exposed to many different types of ASW scenarios in the time that it would take to conduct one scenario in real time.

15.2.1.2 Illustrating Types of Cognitive Readiness in the Context of the *TAO Sandbox*

Because it is a rich and highly configurable environment for providing practice, and one that has been adopted and is being utilized for training complex concepts and

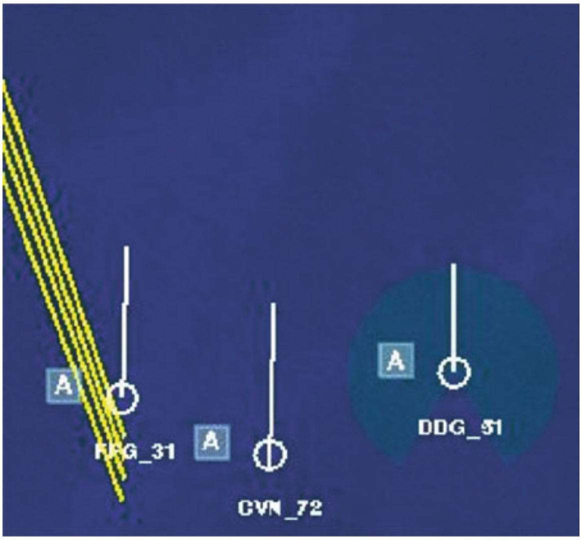


Fig. 15.4 Parallel lines of sound

decision-making, the TAO Sandbox is a good environment for illustrating the types of cognitive readiness training listed above.

Cognitive Readiness for Categorization. Just as a chess player must be able to categorize situations in a game, TAOs must be able to recognize types of tactical situations. When a surface ship is using passive sonar to search for a nearby submarine, the combat information center draws *lines of sound* that pass through the ships current position in the direction of a submarine that the sonar detects. (Decisions about whether to use active or passive sonar are often complex and highly dependent on context. Passive sonar detects by listening for the sounds made by another ship or a submarine; active sonar transmits a loud pinging sound through the water and analyzes the echo of the sound when it bounces back from an object. A much-oversimplified way of thinking about it is to say that passive sonar is more appropriate if you really don't want the submarine to detect your ship. Using active sonar makes your ship detectable at many times the normal range. On the other hand, when the opposing submarine is very silent, passive sonar may not be able to detect the sub even when it is very close, and only active sonar has a reasonable chance of detecting the enemy. At other times, active sonar can be used to "chase away" a submarine that does not want to engage or even to be detected.) Plotting a new line of sound every few minutes using passive sonar produces a *pattern* of lines of sound, when there is a sustained sonar contact. Different patterns imply different types of relative motion relationships between the surface ship and the submarine.

When the lines of sound are parallel to each other, it suggests that the surface ship—in Fig. 15.4, a frigate (a friendly warship, smaller than a destroyer) designated FFG_31—and the submarine are on a converging course. (Each ship has a white bearing line showing its direction of travel. Longer bearing lines reflect higher speeds.) It is important that the TAO recognize examples of this *converging* pattern,

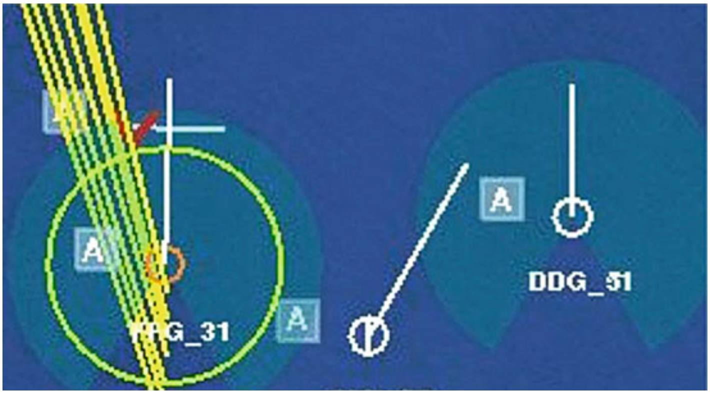


Fig. 15.5 Possible response to converging courses

because it is unlikely that an undetected, silently operating submarine could match speed and course with the frigate. Although the submarine's exact position is not known, it is clear that there is a submarine ahead off the port side (port is the left-hand side of an observer on the ship who is facing the bow) and that it is moving toward the frigate's course. Recognizing this pattern is an example of the first type of cognitive readiness: the ability to categorize objects, events, or conditions in the task context.

Cognitive Readiness to Act on the Basis of a Categorization. In many cases, a learner must learn how to react appropriately when recognition occurs, ordinarily by choosing to carry out an appropriate action or procedure for the detected condition. In the case of a TAO, depending on the current geopolitical situation and the ship's mission, there may be a number of appropriate responses to a categorization such as "possible converging course with hostile submarine."

In Fig. 15.5, the TAO problem-solver has decided on two courses of action. Because his mission calls for the preservation of the MEU, the tactician has turned the carrier, which is labeled CVN_72, to starboard (starboard is the right-hand side of the ship for an onboard observer facing the ship's bow), in order to keep it out of the reach of the submarine's torpedoes. In addition, the TAO has turned on the FFG's active sonar, which has found the exact location of the submarine (shown as a red V shape to the north of the FFG_31 position).

Another pattern that must be recognized is the *passing* pattern, shown in Fig. 15.6. When a series of lines of sound crosses like this, it most likely means that there is a submarine on the port side that is traveling in roughly the opposite direction that the surface ship is. The questions to ask are how near the submarine is and how fast it is traveling.

In this case, active sonar detects the sub quite close to the frigate, just after they have passed each other. Figure 15.7 shows the situation a short while later, with the passing pattern further developed and the position of the submarine detected by the frigate's active sonar.

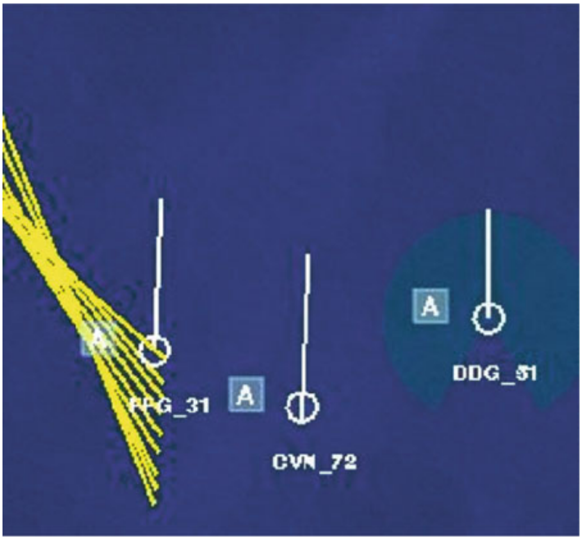


Fig. 15.6 Passing pattern in lines of sound

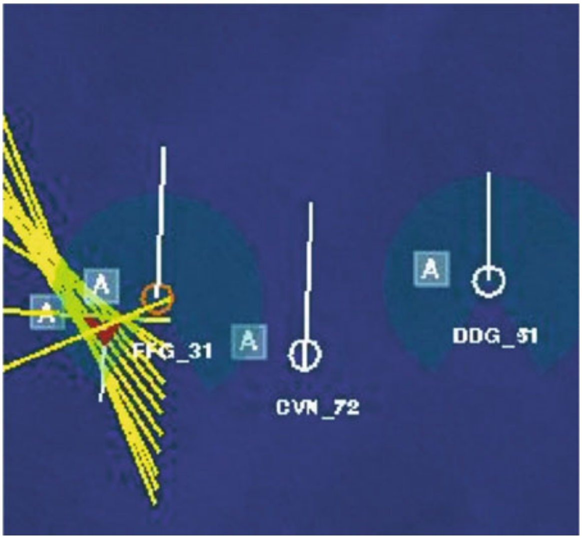


Fig. 15.7 Submarine revealed as it passes

There are many other types of tactical situations that can be categorized in the TAO Sandbox. In addition to many other patterns in lines of sound from passive sonar, there are conditions related to potential danger areas that friendly ships should try to avoid. The emergence of unknown radar contacts with characteristics that mark them as potential threats that should be labeled as *Critical Contacts of Interest* (CCOI) is another.

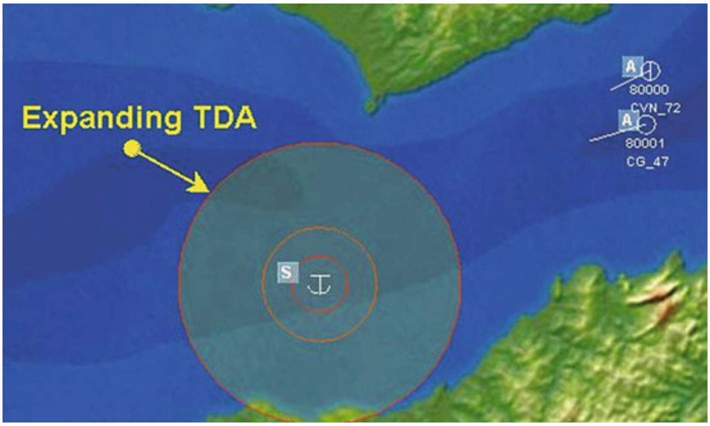


Fig. 15.8 The Torpedo Danger Area (TDA) will expand to fill the strait

In general, training simulations and games can contribute to the training process if they automatically detect opportunities for categorization. In some cases these simulations or games might provide hints or guidance when such conditions are recognized. More generally, however, a simulation or game can inform a collaborating instructional component of such conditions. The instructional component could be a human instructor that is prepared to assess performance or to offer guidance or instruction. Or the instructional component can be a software application that makes analyses based on student actions and emergent conditions, as reported to it by the simulation or game.

Cognitive Readiness to Accommodate Special Conditions in Procedure Execution. Most complex procedures include conditional branching. For example, in ASW, a report of evidence of a hostile submarine at a particular time near a specific point is used to create a point called a datum, with an expanding circle of threat area. Actually, there are two expanding circles: the Furthest On Circle (FOC), which represents the furthest that the now-hidden sub might have traveled at its maximum silent operating speed, and the TDA, which includes the maximum effective attack distance for the torpedoes with which the submarine is presumed to be equipped. In general, TAOs are expected to work to keep friendly ships and especially MEUs out of TDAs.

Sometimes, however, operations may take place in restricted water space that precludes avoidance of the TDA. Normal procedures for avoiding a TDA will not work in such circumstances. Consider the case of the nuclear aircraft carrier CVN 72 and its escort, the guided missile cruiser CG-47. They are to travel west through the Strait of Gibraltar, but there has been a report of a hostile submarine in the strait, originally located at the center of the circle in the figure shown at the left, below. By the time the friendly ships get there, the TDA will have expanded to fill the western end of the strait. Avoidance by maneuvering alone is not possible. See Fig. 15.8.

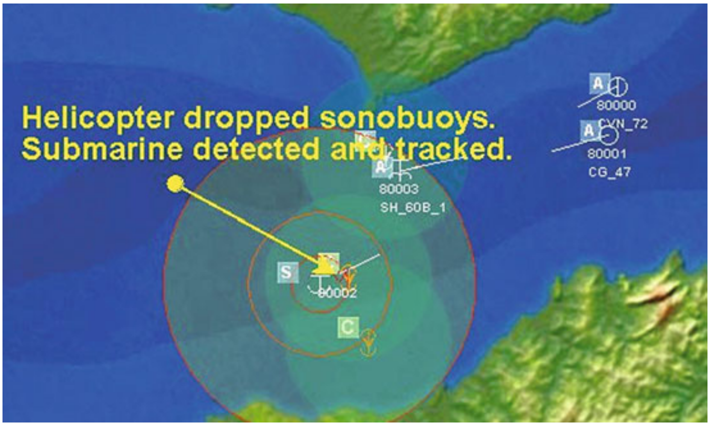


Fig. 15.9 After using sonobuoys to detect the sub, parts of the TDA are safe

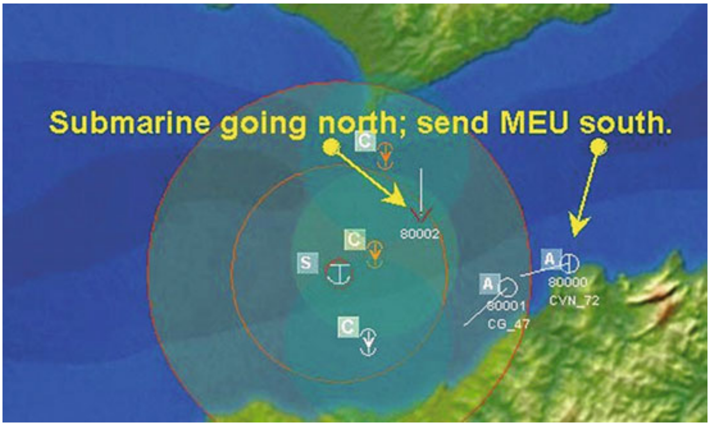


Fig. 15.10 Sending the MEU south to avoid the detected submarine

A solution, in this case, is to launch a helicopter and drop sonobuoys within the TDA to determine the actual position of the submarine. Figure 15.9 shows that two sonobuoys have detected the submarine. The submarine is tracked as moving toward the north-northeast. In this case, the TAO can plan to enter the southern part of the TDA, because the submarine position is known, and it is too far to the north to detect the friendly ships, if they are not moving too fast, and it is too far for its torpedoes to reach the friendly ships. So this can be viewed as an exception to the rule about entering the TDA. See Fig. 15.10.

Alternatively, the TAO can move the datum to the position at which the sub was just detected. The TDA for this new position will be much smaller, because the sonar is showing exactly where the sub is. If the tactician takes this approach, the

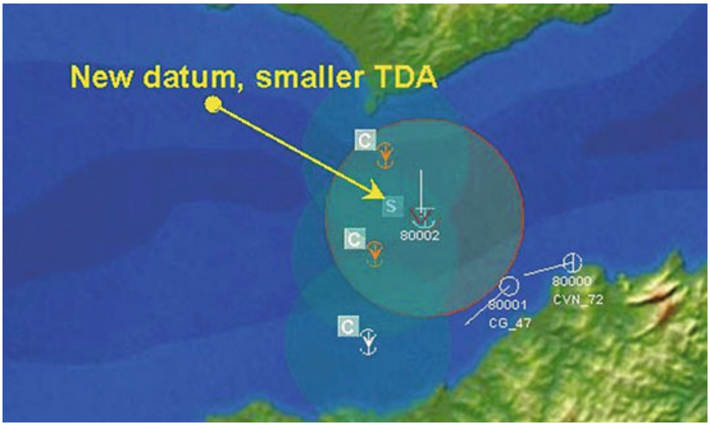


Fig. 15.11 A new datum, a smaller TDA

Fig. 15.12 Dealing with two types of tasks at once

route to the south does not violate the “don’t enter a TDA” rule, because the TDA is now smaller and centered further north, Fig. 15.11.

If the TAO player chooses to either route a path through a “known safe” part of a TDA or chooses to relocate (and thereby resize) the datum and its TDA circle, he or she thereby demonstrates a cognitive readiness to accommodate special conditions in the execution of the “avoid TDA” procedure.

Cognitive Readiness for Task Complexity. There are many possible types of task complexity. For example, someone might have to recognize the need to carry out two procedures at the same time. In Fig. 15.12, the TAO is dealing with attacks from both air and surface units.

Fig. 15.13 Executing two procedures at the same time

Two different procedures need to be conducted at the same time. First, the tactician has to notice and defend against each incoming missile with one or more Surface to Air Missile (SAM) launches. This is a surface unit's major defense against air attack, including surface missile attacks. Second, the source of most of the missile launches, the corvette in the upper left part of the scene, must be eliminated as an ongoing threat. (A corvette is a warship, typically smaller than a frigate.) That requires using a special type of attack, using multiple Harpoon missiles from different ships in a coordinated attack. (The Harpoon is a fairly long-range US ship-to-ship missile.) The defense systems of the corvette will be overwhelmed by the simultaneous strikes from different directions (Fig. 15.13).

A TAO trainee who carries out more than one type of procedure simultaneously shows cognitive readiness for task complexity.

Cognitive Readiness to Generate New Types of Solutions to Problems. Finally, in complex domains such as surface warfare tactics, professional tacticians need to be continually conceiving of new tactical situations and how they could be dealt with. How would ASW tactics change if an opponent's sub could attain higher speeds in subsurface travel than friendly ships can on the surface? If the subs are noisy enough, we could detect them coming, but how would we defend? It is unlikely that brilliant novel solutions to new problems occur in the heat of battle, but expert practitioners may "toy" with such imagined situations in advance of need. They could devise possible approaches to these new solutions and could consult with other expert practitioners about such problems and solutions. A tool like the TAO Sandbox might be able to play a role in encouraging the development of "what-if" scenarios that could eventually be used in real life.

The development of novel situations and solutions can be supported in the Sandbox when users have access to Instructor Mode. In this mode, novel problems can be designed (such as dealing with an attack submarine that is faster than targeted friendly ships). At present, the Sandbox is not instrumented to monitor and report on scenario development. However, once a novel scenario has been designed, it can be tested in Problem Mode, which does report actions and events for evaluation.

15.3 Software Architectures for Assessing Cognitive Readiness

In order to assess cognitive readiness in the context of a game or simulation, there must be an evaluator component that observes performances in that context and makes judgments. The game or simulation will typically need to report three types of information to the evaluator component:

- Initial conditions that may be relevant for assessing performance
- Meaningful actions taken by the player in the game
- Emergent events that may contribute to the evaluator's assessment of the player

Assessment may be conducted naturalistically, by giving the user of the simulation (or the player of a game) a problem and then making judgments based on the actions and events that occur during that problem session. Alternatively, more artificial or staged assessments can be conducted, by presenting a particular problem situation, perhaps with an event history that brought the user to this point, and then offering a menu of possible actions to the user and judging the menu choice made. This approach would be more like conventional assessment and would permit a more controlled presentation of "test items." This chapter deals with the more challenging design and implementation of "naturalistic" assessments in the context of valid problems or scenarios.

A software architecture like the one sketched below would be an appropriate one for conducting such naturalistic assessments in the context of game scenarios or simulations (Fig. 15.14).

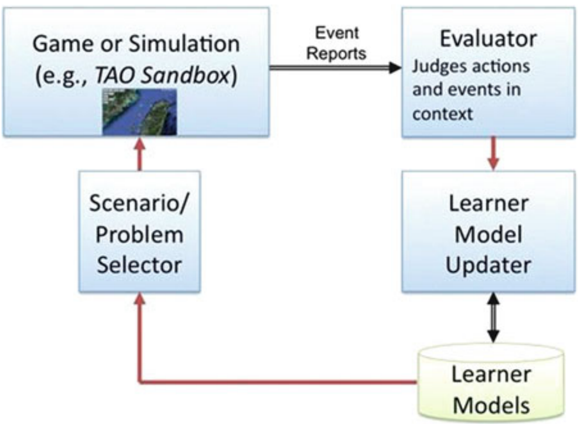


Fig. 15.14 Architecture for assessment in simulation contexts

The evaluator object uses reports of events (and relevant initial conditions) to make judgments about what the player of the game knows about the domain and to evaluate the user's play. These judgments, which could be quite detailed—that is, they could reflect detailed elements of domain knowledge—are passed to an object that is responsible for updating the model of the learner. Learner models are stored in a database. The Learner Model Updater needs to be able to read the model for the current student, because decisions about the goodness of certain actions may be based, in part, on an assessment of how sophisticated the learner already is in the domain. When it is time to select a new scenario or problem in order to continue assessing the learner, the learner models database is accessed, and the selector chooses a problem that will reveal some aspect of learner knowledge that has not already been adequately tested in previously presented scenarios.

Note that this architecture does not specify *how* the next problem or scenario is to be selected. That is a matter for particular implementations of the architecture. One approach would be to store meta-data with problems that reflect the concepts and skills required to perform well on the problem. This map of problem-relevant knowledge could be compared with the current learner model for a particular student, and the problem chosen could be the one which utilizes mostly understood concepts or skills in evidence, but that also requires an increment of additional knowledge or skill that is not yet attested in the learner model.

The evaluator in this architecture must have detailed information about how to judge the goodness of actions and events in particular contexts. One way to achieve this is to code a special-purpose evaluator, adhering to a module interface standard, for each game or scenario that must be evaluated. Another way would be to use a generic evaluator, but to import a set of evaluation rules for the specific domain. These rules could be a set of conjoint Boolean expressions, specifying, for example, that if threat values are low, then entering a TDA is OK, but, if threat values are high, then entering a TDA, especially for the MEU, is bad.

Evaluator judgments can serve to stimulate instructional interactions either during a scenario-based practice session or during an after-action review (AAR).

An expanded version of this architecture for evaluating in game contexts can also be utilized to teach in the context of the game or simulation, as shown in Fig. 15.15.

In this case, some of the evaluator's judgments are also passed directly to an interactive instruction Control object, so that the player can have obvious mistakes pointed out immediately. In addition, however, the learner model updater's synthesis of current evaluations with the learner model can produce insights that can be conveyed under the control of the Instruction Control object. In the above diagram, there is an optional element called the instruction/explanation interface. Some games and simulations might provide interface elements for presenting newly generated explanatory text in the game context. The TAO Sandbox can present very brief text messages. But many games will not have such a subsystem, so a separate module should be available for instructional presentations, especially longer ones. Note that the explanation interface would not have to be textual: text to speech, prerecorded voice, videos, or animations could be presented under the control of the interactive instruction component. Different implementations of the instruction/explanation interface might be utilized for different media types.

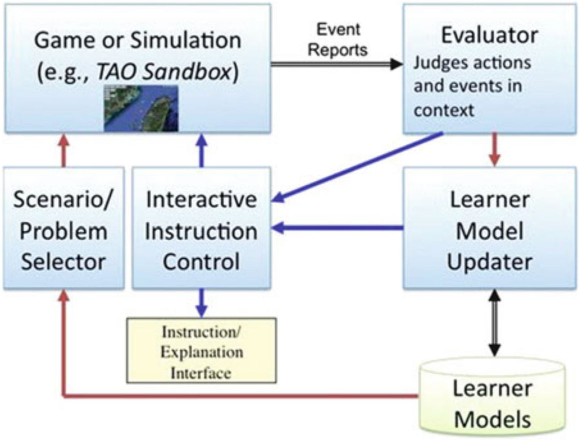


Fig. 15.15 Architecture for assessing and instructing in simulation contexts

15.4 Software Support for Assessing Cognitive Readiness

A game or simulation can support assessment of cognitive readiness by a collaborating application that evaluates student actions in context, such as the evaluator object in the above figures. To be practical, the game must report at a level that is meaningful in terms of the game-player's tasks. That is, if the game only reported that, for example, "the left mouse button was clicked at the point [293, 732]," the action evaluation component would have to maintain its own internal simulation of the game domain in order to make sense of such low-level information. Examples of meaningful reports in the TAO Sandbox would have meanings such as "player launched a SAM from CG-47 against surface-to-air-missile-16 at 14:22:05" or "player turned CVN-72 to bearing 30 at 16:05:00." Here is an actual example of a report of a user action:

```
00:07:29 .k.s_CG_47. LoadandLaunchSH_60B CG_47
```

This means that at the virtual game time of 7 min and 29 s, the cruiser CG-47 loaded and launched an SH-60B helicopter, a helicopter equipped to assist in ASW. (Note: the ".k.s." is part of the simulation's designation of the cruiser CG_47. It can be ignored). The next message from the Sandbox provided specific information about where the helicopter appeared, and what its initial heading and speed were:

```
00:07:29 .k.helo_master. LaunchHelo SH_60B3 400 353 324 .k.s_CG_47.
```

This means that the SH-60B helicopter was launched at the location [400, 353] in the coordinate system used by the Sandbox and that its bearing was 324°.

It is not enough, however, to simply report player actions. It is also necessary that the game must detect and report the occurrence of instructionally relevant emergent events. For example, the TAO Sandbox has semi-intelligent hostile submarines. When they are assigned hostile behavior, they actively search for surface ships using their passive sonar. When they detect one, they move toward it until they come within visual range. Then they briefly raise their periscopes to see whether the sonar hit is actually a desirable target. If it is, they submerge again and close to torpedo range and then fire. This autonomous, rule-based behavior interacts with the results of actions previously taken by the game player. Torpedo attacks cannot be precisely scripted in advance, because they will only occur if the player maneuvers in such a way that the submarine can detect and close on a friendly ship position. Such *emergent* events must be reported by the game to the analyzing evaluator or that component will be unaware of them. There are many examples of such emergent events that the evaluator might need to know about, including a friendly ship entering a TDA, a friendly ship coming within torpedo range of a hostile sub, and a torpedo being fired by a hostile sub. Without being told that a torpedo was launched, the evaluator would have no idea why the targeted friendly ship was taken to flank speed by the user and put through a rapid succession of bearing changes. Here is an example of the report of an emergent event:

```
00:59:45 .k._happenings. friendly_in_TDA CVN_72
```

When the simulation detected that one of the user's ships (a carrier named CVN_ 72) entered a TDA, it announced to the evaluator component that there was a "friendly_in_TDA" event. (This announcement was made by the simulation's .k._happenings object.)

At almost one (virtual) hour into the session, the carrier CVN-72 entered a TDA. Another type of emergent event occurred just over 15 min later, when the carrier was torpedoed:

```
01:15:22 .k._happenings. friendly_killed CVN_72
```

This time the .k._happenings simulation object announced that there was a "friendly_killed" event involving the carrier CVN_72.

A third type of information that a game must provide to an evaluator component is aspects of the initial state of a game session. For example, if a submarine from a country with which the friendly force is at peace is the source of a datum, then that TDA probably does not actually represent a danger area. In this circumstance, it may be appropriate not to avoid the TDA, but to simply go through it on the way to one's appointed task. Initial threat levels and other aspects of the initial state of the game must be reported. In the TAO Sandbox, it is the instructor who authors a scenario who determines what threat level a student should expect. The author writes a mission briefing that describes the situation in sufficient detail that the user should be able to estimate whether there is no threat at all (level 0) or a very high threat level (10), representing open war with a competent opposing force. Because the Sandbox cannot read and understand the briefing, the author must give the Sandbox a clue about the threat by specifying a number in the range 0–10. At run-time, the Sandbox will tell the evaluator about this threat, so that it can take the threat level into account in judging the users actions. (For example, going through elaborate maneuvers that delay accomplishing a mission in order to avoid entering sensor range of an allied submarine—a low threat situation—would not make sense. Going through the same maneuvers in a time of war, to avoid a hostile submarine, might be the correct course of action. The evaluator component has to know the threat level to accurately evaluate actions in context.) In a sense, such initial reports are a special case of emergent events: they are all the events that emerge the instant the game session starts. Here is an example of such an initial state element:

```
00:00:00 Threat_Level 0
```

This means that this scenario is one that, at least initially, is conducted in an environment that is one of relative peace and security. The evaluator will not expect TDA avoidance maneuvers to be carried out in such a scenario.

15.5 Software Support for Teaching Cognitive Readiness

Many of the low-level services described by Munro (2007) that a game or simulation should provide for training will be required by an interactive instruction controller that aims to provide a level of cognitive readiness for specific tasks. These low-level services include these:

Highlight object
Unhighlight object
Set game internal value
Emulate user action
Register interest in an occurrence
Report occurrence

An instructional component can provide remediation or explanations in the simulation context by utilizing some of these services. A request to highlight an object is interpreted by the Sandbox as a direction to create a large semitransparent arrow pointing at the object. Comment objects, which are textual elements that can be made to appear in the simulation window, can be given text values, such as "Using active sonar when trying to avoid a hostile submarine is dangerous," if the instructional component wants to present that message to the simulation user.

Let us consider the five types of cognitive readiness that we listed at the beginning of this essay, looking at them in the context of the TAO Sandbox.

- *Cognitive Readiness for Categorization*. To teach about categories of situations, it is necessary to draw attention to salient characteristics of examples. One way to do that is to highlight graphical elements that illustrate those characteristics. Setting the internal value that names the file with the current scenario data causes that file to be read into the TAO Sandbox.
- *Cognitive Readiness to Act on the Basis of a Categorization*. In addition to using the highlight and unhighlight features, in order to teach a user what to do after recognizing a salient condition, it may be necessary to demonstrate a procedure under control of the instruction module. That requires the use of the *emulate user action* service.

- *Cognitive Readiness to Accommodate Special Conditions in Procedure Execution*. In order to teach learners how to respond to special conditions, the instructional control module needs to recognize when those conditions arise. One way to do that is to register an interest in those conditions and then to receive reports of their occurrence. To teach how to respond to those specific situations, emulating user actions may again be required.
- *Cognitive Readiness to Categorize Complex Situations and Act Appropriately, Prioritizing or Combining Procedures*. Teaching users how to recognize simultaneous conditions and the need to carry out combined procedures or procedures done in parallel requires providing many practice examples, using highlighting, explanations, and demonstrations, as required. The work of Sweller and Cooper (1985) showed that worked examples contribute to effective learning when combined with opportunities to solve problems. Scenario demonstrations play the role of worked examples in TAO Sandbox training.
- *Cognitive Readiness to Generate Solutions for Novel Problems*. Teaching people how to generate new types of solutions to novel problems is the area in which we are least prepared to offer specific software support. The problem is that, if the solution really is novel, it will take an extremely intelligent software system to come up with useful novel solutions on its own. (If this is possible, perhaps the software should be solving the problems, rather than training human beings to do so!) Less ambitiously, the training system might try to evaluate the goodness of a novel solution. But if the novel solution actually violates some of the standard metrics of "goodness," even this could be quite difficult to do. Novel solutions are most likely to be achieved by expert, motivated, professional practitioners who are willing to "play" with solutions to hypothetical problems. Perhaps the best software support that we can offer for this level of cognitive readiness is easy-to-use practice environments that promote trying alternatives and sharing interesting outcomes with others.

15.6 Implementation

The TAO Sandbox is a simulation that has been implemented and continues to evolve as it is used to develop and deliver interactive scenarios for students at SWOS. Instructors and the professional in-house staff of the school utilize the Sandbox in authoring mode to build tactical planning problems in a variety of simulated contexts. This simulation has been instrumented to report actions and emergent events to a collaborating evaluator component. Ongoing collaboration with the designers and developers of one such component, the CRESST Assessment Application (CAA) is currently underway. The Sandbox is undergoing additional instrumentation in order to report to the CAA on the goodness of pedagogically relevant actions and events. The CAA will use a Bayes Net to summarize high-level estimations of the competence of the problem-solvers.

The TAO Sandbox also offers services to support instruction. Working with our CRESST colleagues, we are developing approaches to generating and delivering instructional interactions in the context of the Sandbox. These features will constitute the interactive instruction component shown in the architecture diagrams above.

15.7 Conclusions

Five levels of cognitive readiness for performing tasks have been described. Each of these levels has been illustrated in the context of a game-like tactical planning tool, the TAO Sandbox, which is used for conceptual training of TAOs. A simple architecture for assessing cognitive readiness in the context of games and simulation objects such as the TAO Sandbox was presented, together with a somewhat more complex architecture for teaching in such contexts. Games and simulations that are to work with assessor/evaluator components in a training system must report instructionally relevant elements of initial problem states, meaningful actions taken by users, and certain potentially relevant emergent events. Instruction in such contexts requires that the game or simulations offer certain low-level services to an interactive instructional component. When games and simulations that are to be used to help people learn to perform tasks are developed with such services, their potential to play a useful role in advanced training systems is much greater than it would otherwise be.

Acknowledgements The work reported herein was partially supported by a grant from the Office of Naval Research, Award Number N00014-08-1-0126. The findings and opinions expressed here do not necessarily reflect the positions or policies of the Office of Naval Research.

References

Auslander, B., Molineaux, M., Aha, D. W., Munro, A., & Pizzini, Q. (2009). *Towards research on goal reasoning with the TAO Sandbox* (Technical Note AIC-09-155). Washington, DC: Naval Research Laboratory, Navy Center for Applied Research on Artificial Intelligence.

Clark, R. E., Yates, K., Early, S., & Moulton, K. (2010). An analysis of the failure of electronic media and discovery-based learning: Evidence for the performance benefits of guided training methods [electronic version]. In K. H. Silber & R. Foshay (Eds.), *Handbook of training and improving workplace performance, Volume I: Instructional design and training delivery*. Washington, DC: International Society for Performance Improvement.

Endsley, M. R. (1995). Toward a theory of situation awareness in dynamic systems. *Human Factors, 37*, 32–64.

Mayer, R. (2004). Should there be a three-strikes rule against pure discovery learning? The case for guided methods of instruction. *The American Psychologist, 59*, 14–19.

Merrill, M. D. (2002). First principles of instruction. *Educational Technology Research and Development, 50*(3), 42–59.

Munro, A. (2007). Foundations for software support of instruction in game contexts. In H. F. O'Neil & R. S. Perez (Eds.), *Computer games and team and individual learning*. Amsterdam: Elsevier.

Munro, A., & Pizzini, Q. A. (2009). *The TAO Sandbox user guide*. Working paper, Center for Cognitive Technology, University of Southern California.

Munro, A., Pizzini, Q. A., & Bewley, W. (2009). Learning anti-submarine warfare in the context of a game-like tactical planner. In *Proceedings of the Interservice/Industry Training, Simulation & Education Conference (I/ITSEC)*. Retrieved from http://ntsa.metapress.com/ling.asp?id=u1q691546361167

Paas, F., Renkl, A., & Sweller, J. (2004). Cognitive load theory: Instructional implications of the interaction between information structures and cognitive architecture. *Instructional Science, 32*, 1–8. Retrieved from http://meds.queensu.ca/ohse/assets/cogntive_load_theory_pass_2004_.pdf

Sweller, J., & Cooper, G. A. (1985). The use of worked examples as a substitute for problem solving in learning algebra. *Cognition and Instruction, 2*(1), 59–89.

Chapter 16
Cognitive Readiness for Complex Team Performance

Clint Bowers and Janis Cannon-Bowers

16.1 Introduction

According to Morrison and Fletcher (2002), cognitive readiness can be described as "the mental preparation an individual needs to establish and sustain competent performance in the complex and unpredictable environment of modern military operations" (p. 2). These authors go on to posit that the relevant components of cognitive readiness are "situation awareness, memory, transfer of training, metacognition, automaticity, problem-solving, decision-making, flexibility and creativity, leadership, and emotion" (p. 3). They argue that these components of cognitive readiness are both measureable and amenable to training. Hence, they recommend that cognitive readiness be considered an important precursor to successful operational performance.

While not explicit in the above definition, many have argued that the most complex and taxing performance demands in typical operational environments are often associated with the requirement to function as part of a team (Cannon-Bowers & Bowers, 2010; Fletcher & Wind, this volume). According to Cannon-Bowers and Bowers (2010), most high-performance environments present demands on individual performers that transcend their individual tasks. These demands have often been labeled teamwork demands (Morgan, Salas, & Glickman, 1993) as opposed to taskwork demands. This line of thinking implies that to be cognitively ready to confront a complex environment, individuals must hold at least some level of competence for teamwork.

C. Bowers (✉)
Department of Psychology, University of Central Florida, 4000 Central Florida Boulevard, Orlando, FL 32816-1390, USA
e-mail: bowers@mail.ucf.edu

J. Cannon-Bowers
Institute for Simulation & Training, University of Central Florida, 3100 Technology Parkway, Orlando, FL 32826, USA
e-mail: jcannonb@health.usf.edu

H.F. O'Neil et al. (eds.), *Teaching and Measuring Cognitive Readiness*,
DOI 10.1007/978-1-4614-7579-8_16, © Springer Science+Business Media New York 2014

To expound on the notion that cognitive readiness includes a teamwork component, this chapter provides an initial definition and explanation for some competencies, which we believe are components of cognitive readiness, that specifically address the ability to perform as part of a team. First, we provide a simple model of cognitive readiness for team performance, and then we define a set of competencies that we believe are most crucial as precursors to effective team performance. Next, we describe a set of approaches for measuring cognitive readiness skills that contribute to teamwork. Finally, we discuss a set of interventions that can help achieve a good match between cognitive readiness skills and associated teamwork demands.

16.2 Cognitive Readiness Demands for Teamwork

Figure 16.1 displays a model of cognitive readiness as it is related to team performance. According to Fig. 16.1, individual cognitive readiness factors are hypothesized to affect individual performance (left side of the figure). This is the traditional way in which the construct of cognitive readiness has been assumed to operate. Our contention is that a subset of cognitive readiness factors is also associated with effective team performance (upper path in Fig. 16.1). Borrowing from Morris and Fletcher's (2002) definition, these individual cognitive readiness factors can be considered necessary to mentally prepare the individual to establish and maintain performance in a complex *team* environment.

Inspection of Fig. 16.1 also reveals that we believe that the combined influences of both individual cognitive readiness and cognitive readiness for teamwork are responsible for observed performance. In other words, individuals must be competent in the taskwork associated with their individual jobs and also proficient in the competencies required to be an effective team member. This interpretation is highly consistent with the prevailing view among team scholars that overall performance is best described

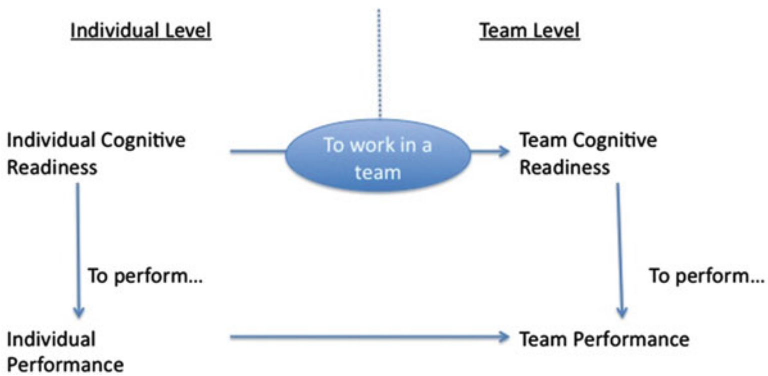


Fig. 16.1 Cognitive readiness and team performance

as being influenced by both individual and team-level competence (Cannon-Bowers & Bowers, 2010). The following sections describe some of the components of cognitive readiness that we believe are necessary for team performance.

16.3 Individual Cognitive Readiness Skills Required for Effective Team Performance

As noted, several team researchers have offered taxonomies of team-level competencies that are hypothesized to underlie effective team performance (Stevens & Campion, 1999; also see Cannon-Bowers & Bowers, 2010, for a review of various taxonomies). Typically, the term "competency" has been used in these cases to refer to the knowledge, skills, and attitudes (KSAs)—or knowledge, skills, and abilities (also KSAs), depending on the authors' position—that are unique to team situations (we prefer to use the "A" to refer to attitudes). Further, several researchers have argued that the nature of the task and team situation will affect which team-level competencies are needed and how they are likely to be expressed within the team context (Cannon-Bowers, Salas, Tannenbaum, & Mathieu, 1995).

While it is often assumed that teamwork competencies can be developed through experience in the team situation itself, our contention here is that it is prudent to think about which teamwork competencies may actually be considered prerequisites to effective team membership. In this sense, we believe that at least some of the team competencies hypothesized in past work can be considered as part of cognitive readiness for team performance. Importantly, the implication of this line of thought is that *cognitive readiness for team performance should be achieved* (*and assessed*) *prior to the individual joining the team.*

The major components of cognitive readiness for team performance are summarized in Table 16.1 and described in the following sections. This is not intended to be an exhaustive list; rather it represents what we believe are the most important and potentially useful components. We address the measurement and intervention issues in subsequent sections.

16.3.1 Knowledge of Teamwork

The first component of cognitive readiness for teamwork shown in Table 16.1 is *knowledge of teamwork*. It is frequently assumed that all workers possess the knowledge required for teamwork, and that no additional preparatory training is required. However, team performance researchers have questioned this assumption (Salas, Prince, et al., 1999; Stevens & Campion, 1999). Indeed, many have argued that team members must possess a degree of declarative knowledge about how to be a team member as a prerequisite to developing effective teamwork skills (Cannon-Bowers et al., 1995; Stevens & Campion, 1997). According to Cooke et al. (2003),

Table 16.1 Cognitive readiness competencies for teamwork

Competency	Definition	Measurement approach	Intervention
Knowledge of teamwork	Declarative knowledge about how to be a team member (Cannon-Bowers et al., 1995)	Situated judgment tests; subjective scales	Traditional classroom training; scenario-based training
Leadership	KSAs required to organize and direct a team's work (Burke et al., 2006)	Situated judgment tests; behavioral observation	Selection; scenario-based training
Mutual performance monitoring/back-up	"The discretionary provision of resources and task-related effort to another member of one's team that is intended to help that team member obtain the goals as defined by his or her role" (Porter et al., 2003, p. 4)	Behavioral observation	Cross-training
Communication	The timely and clear provision of information (Bowers, Braun, & Morgan, 1997)	Behavioral observation	Selection; classroom training; scenario-based training
Interpersonal skills	A set of skills including conflict management, motivation, and managing affect (Mathieu, Maynard, Rapp, & Gilson, 2008)	Situated judgment tests	Selection; composition; skill training
Positive teamwork attitudes	Degree to which workers value or prefer working in a team (Helmreich & Foushee, 1993)	Subjective scales	Selection; composition; classroom training

empirical research appears to support this contention (as cited in Hirschfeld, Jordan, Feild, Giles, & Armenakis, 2006). For example, Cooke et al. (2003) demonstrated that not only knowledge of teamwork predicted short-term task performance but also new knowledge obtained in earlier trials predicted performance in subsequent trials. Likewise, Stevens and Campion (1999) have shown that knowledge of teamwork is a predictor of effective team performance.

Taken together, these findings suggest that building appropriate knowledge of teamwork is a worthwhile endeavor before allowing individuals to participate as team members. In this sense, knowledge of teamwork can be considered a component of cognitive readiness for teamwork.

16.3.2 Leadership Skills

Several authors have discussed the importance of leadership behaviors in team performance (Bass, 1990; Hackman & Wageman, 2005). Given that the importance of team leadership is generally accepted, there has been relatively little attention paid to the nature of the specific behaviors that individuals must perform so that effective leadership can occur (Kozlowski, Gully, Salas, & Cannon-Bowers, 1996). However, researchers have recently attempted to address this question.

For example, Burke et al. (2006) published a meta-analysis designed to identify individual leadership skills necessary for teamwork. Boundary spanning was one of the skills that accounted for a great deal of variance in performance. *Boundary spanning* refers to a collection of political skills and abilities that increase a team's resources and the information available to them (Brown & Eisenhardt, 1995). Interestingly, empowerment, or the ability to encourage team members to manage their own behavior, was also strongly related to team performance. This behavior seems to be important in creating a culture of continuous learning that allows the team to analyze and adjust its behaviors to optimize its performance.

Hence, at least a few specific skills associated with effective leadership have been identified as possible components of cognitive readiness for team performance. More research along these lines is needed to better refine and expand these concepts.

16.3.3 Mutual Performance Monitoring/Back-up Behavior

There is little doubt that the requirement to coordinate individual behaviors is an essential aspect of team performance. One example of effective coordination is when one team member "backs up" a colleague. *Back-up behavior* has been defined as "the discretionary provision of resources and task-related effort to another member of one's team that is intended to help that team member obtain the goals as defined by his or her role" (Porter et al., 2003, p. 391).

Teams are often used in high-risk, complex environments simply to allow this kind of back-up behavior. However, some researchers have expressed concerns about the wisdom of encouraging back-up behaviors without restraint (Bowers et al., 1997). They point out that the cognitive behaviors required for back-up include monitoring, communication, coordination, temporal planning, and so forth (Barnes et al., 2008; Burke, Weir, & Duncan, 1976). Further, Bowers et al. (1997)

argued that these behaviors all require cognitive resources, and they hypothesized that they may actually hurt team performance during times of high workload.

The empirical literature regarding back-up behavior is mixed. For example, several studies have demonstrated that back-up behaviors by team members are positively related to team performance (Porter, 2005; Porter et al., 2003). However, Barnes et al. (2008) demonstrated that back-up behaviors may place teams at risk to overlook other critical tasks, and therefore must be deployed with caution. Furthermore, the provision of back-up behaviors was related to subsequent decreases in effort by the member receiving the help. Therefore, one might define the competency of back-up to include not only the ability to help a teammate but also the ability to make a decision about when sufficient cognitive resources are available to provide such assistance. It also seems to require an ability to receive help without diminishing subsequent effort. Finally, DeChurch and Haas (2008) have pointed out the importance of anticipating possible back-up situations *before* task performance.

Taken together, these findings indicate that the seemingly straightforward notion of backing up a teammate when he/she is in need is actually a fairly complex behavior that requires careful thought and preplanning. Hence, it appears that back-up behavior has a sizable cognitive component. While some of this type of hands-on practice is probably essential to fully develop this skill in the team environment, we contend that viewing back-up behavior as part of cognitive readiness for team performance implies that the underpinnings of this skill can be developed as a precursor to team performance.

16.3.4 Communication Skills

Communication skills are included in almost every description of the factors that contribute to effective teamwork. This is likely due to the fact that communication is the primary manner by which teams share information and resources. A variety of errors and other negative outcomes have been associated with breakdowns in team communications (Cannon-Bowers & Salas, 1998a; Christensen et al., 2000; Wears et al., 2003). Communication is a foundation of each of the transition processes that Marks, Mathieu, and Zaccaro (2001) have hypothesized to be crucial to team performance. These processes include planning, goal specification, and strategy formulation. A recent meta-analysis by Mesmer-Magnus and DeChurch (2009) supports this contention.

Despite the fact that there is widespread agreement that communication is important, there is substantially less clarity about what elements of communication are reliably associated with effective team performance. One approach to understanding this phenomenon has been to analyze the frequency of overall communication, or of specific subtypes of communication contents. For example, Patrashkova-Volzdoska, McComb, Green, and Compton (2003) failed to support the hypothesis that increased communication would be related to better team performance. Rather, they found a curvilinear relationship, with both excessive and insufficient communications being related to poorer performance. Similarly, MacMillan, Entin, and Serfaty (2004) have described the deleterious effects of excessive communication.

In addition to communicating in correct amounts, there are some data to suggest that *how* team members communicate is more important than how much. For example, it has been demonstrated that effective team members ask direct questions rather than implied ones (Urban, Bowers, Monday, & Morgan, 1995). Similarly, meta-communication skills also seem related to team performance (Svensson & Andersson, 2006). It has been suggested that the ability to communicate in "closed-loop" patterns (completing one topic before starting another) is associated with effective team performance (Bowers et al., 1997). Researchers also include nonverbal communications among the important communication behaviors in team performance (Fowlkes, Lane, Salas, Franz, & Oser, 1994; Harris & Sherblom, 2002).

Given all that has been said, it is clear that individual communication competencies are associated with subsequent team performance. However, there is not much clarity about which specific aspects of communication competence are most crucial. Based on the available data, one might hypothesize that some of these communication competencies involve the ability to: split attention, maintain motivation after receiving assistance, maintain items in memory until other items have been "closed out," and communicate nonverbally.

Clearly, more research is needed to study the specific impact of these potentially critical skills. In the meantime, we feel justified in including communication skills as part of cognitive readiness for team performance. As with back-up behavior, we cannot argue with the notion that direct practice of communication skills in a realistic team setting is necessary for such skills to be honed. However, we also believe that the foundation for such skills can be laid prior to entering the team setting. What must be determined is exactly which aspects of the communication skill are best developed as part of cognitive readiness for team performance, and which are best trained in that team environment itself.

16.3.5 Interpersonal Skills

Teamwork, by its very nature, requires members to interact with each other, often during times of stress. This requirement has led some researchers to suggest that social (or interpersonal) skills are more important for team performance than they are in non-team workplaces (Campion, Medsker, & Higgs, 1993; Mohrman & Cohen, 1995). As a construct, *interpersonal skills* typically encompass several specific behaviors, such as conflict management, motivating others, and managing affect (Mathieu et al., 2008). It has been demonstrated that a team's interpersonal processes predict its performance (Mathieu & Shulze, 2006), and the satisfaction of its members (Maynard, Mathieu, Marsh, & Ruddy, 2007). Similarly, failures of interpersonal skills have been demonstrated to disrupt a team's processes in the long term (Lingard et al., 2004).

To the degree that individual interpersonal skills are an important determinant of subsequent positive interpersonal processes in teams, they can be considered a requisite factor for team performance. While this link needs to be established more convincingly through additional study, we contend that interpersonal skills are an important part of cognitive readiness for team performance. In this case, much preparation can be accomplished prior to entering the team situation, since there is reason to believe that interpersonal skills are fairly generic with respect to the task and team (Cannon-Bowers et al., 1995). In other words, learning effective interpersonal skills in one setting (i.e., task or team) likely generalizes to other tasks and teams.

16.3.6 Positive Teamwork Attitudes

Prior research has demonstrated that a person's *attitudes towards teamwork* influence the effectiveness with which he/she applies teamwork skills (Helmreich & Foushee, 1993; Helmreich & Merritt, 1998). For the most part, this research has focused on the degree to which people value teamwork or prefer to work in teams. There are empirical data to suggest a link between positive teamwork attitudes and performance. For example, Morey et al. (2002) demonstrated that an attitude-based team coordination training program was effective in improving attitudes and in reducing errors in a medical setting. Helmreich and his colleagues have also described a long line of research that indicates the effectiveness of this type of training in commercial aviation settings (Helmreich & Foushee, 1993; Weaver et al., 2010).

Some researchers have suggested that attitudes towards teamwork are better considered as a multidimensional construct (Alavi & McCormick, 2007) and that each dimension might have a specific intervention for improvement. For example, *teamwork self-efficacy* (or *collective efficacy*) is one factor of teamwork attitudes that has received some attention (Cannon-Bowers et al., 1995; Guzzo & Dickson, 1996). Teamwork self-efficacy is the degree to which an individual believes he/she possesses the skills required to be an effective team member (McClough & Rogelberg, 2003). However, in analyzing the construct, McClough and Rogelberg found that teamwork self-efficacy was not a significant mediator of the relationship between teamwork KSAs and team performance. On the other hand, Stone and Bailey (2007) have recently described a subtype of team self-efficacy, team *conflict* self-efficacy, that specifically emphasizes one's confidence in his/her ability to manage conflict within a team. In validating their scale, Stone and Bailey demonstrated that team conflict self-efficacy was a significant mediator between other team characteristics and subsequent team performance.

It is important to note that research on attitudes towards teamwork and team performance indicates that it may be a complex relationship. For example, some researchers have hypothesized that successful teams will have higher attitudes towards teamwork than unsuccessful ones (i.e., that attitudes are a consequence of performance rather than the other way around). However, a recent study by Wallin, Meurling, Hedman, Hedegård, and Felländer-Tsai (2007) indicates that teamwork attitudes did not improve following a simulation-based training experience, even though team performance did improve. On the other hand, mental attitudes towards teamwork were relatively positive to start with in this study, suggesting that a threshold level might be necessary for effective team performance.

While further research is needed to better understand the phenomenon, it is probably safe to conclude that positive teamwork attitudes are a necessary component of cognitive readiness for team performance. Moreover, this construct likely includes strictly attitudinal aspects (i.e., positive vs. negative feelings about being on a team) as well as more self-referenced aspects (i.e., collective efficacy). These variables should be pursued as a subset of those that can be developed as a precursor for team membership.

16.4 Measuring Cognitive Readiness Skills Required for Team Performance

Referring again to Table 16.1, the third column indicates the type of measurement approach that we believe is most appropriate for each component in question. These approaches are described in more detail in the following sections.

16.4.1 Subjective Tests

Subjective paper-and-pencil tests have been the most common approach to measuring attitudes towards teamwork. Perhaps the best known example of such measures is the Cockpit Management Attitudes Questionnaire (CMAQ, Helmreich, 1984). The CMAQ was originally designed to measure teamwork attitudes in aviation crews, but has since been adapted for use in a variety of settings including medicine (Flin, Fletcher, McGeorge, Sutherland, & Patey, 2003; Helmreich, Sexton, & Merritt, 1997) and the oil and gas industry (Flin, O'Connor, & Mearns, 2002). A factor analytic study by Gregorich, Helmreich, and Wilhelm (1990) demonstrated that there are three reliable clusters of items in the CMAQ: Command Responsibility, Communication, and Recognition of Stressor Effects. Analyses of variants of the CMAQ have yielded other clusters that might be of interest (cf. Crichton, 2005; Flin et al., 2003).

Survey approaches are the predominant technique used to measure other attitudinal factors such as collective efficacy (Alavi & McCormick, 2004). Collective efficacy refers to the average judgment of members of a group that the group can accomplish its chosen task. A variety of measurement approaches for this construct have been used and validated by researchers (Stajkovic, Lee, & Nyberg, 2009; Stout, Driskell, & Salas, 1997).

Finally, there are a few multidimensional scales that address one or more of the competencies previously described. For example, Pulakos, Arad, Donovan, and Plamondon (2000) developed the Job Adaptability Inventory (JAI) and Ployhart and Bliese (2006) developed the I-ADAPT measure. These assessments were not specifically created to predict effectiveness in a team, but several of the items refer to team settings and appear to assess some of the competencies we've listed above. In fact, Pulakos, Dorsey, and White (2006) have recently described why they think the JAI would be an effective predictor of adaptive team performance. However, we are unaware of an empirical evaluation of these hypotheses.

Hertel, Konradt, and Voss (2006) developed a scale to assess competencies to participate in a virtual team, the Virtual Team Competency Inventory (VCTI). These researchers extracted competencies for traditional teams from literature in the area, and then added theoretically derived competencies for virtual team members. The final scale comprised 132 items that were rated using a Likert scale. Although the scale is relatively new, early empirical data support the predictive validity of the inventory (Hertel et al., 2006).

16.4.2 Situated Judgment Tests

Situated judgment is, perhaps, the most popular approach to assess the teamwork knowledge and skills of prospective team members. The Teamwork KSA Test (Stevens & Campion, 1997) appears to be the most frequently used test. This test presents prospective team members with a short vignette and asks them about the appropriate course of behavior. For example, the participant reads a vignette such as "Your team wants to improve quality and flow of the conversations among their members. You should: …" The respondent then chooses among options such as "set up a specific order for everyone to speak and follow it" or "use comments that build upon and connect to what other members have already said."

Empirical studies indicate that the Teamwork KSA Test is effective in predicting subsequent teamwork behaviors and team performance (Stevens & Campion, 1999). Further, it seems that teamwork knowledge adds incremental predictive validity to other team selection approaches, such as personality characteristics (Morgeson, Reider, & Campion, 2005). However, Miller (2001) indicated that the test may be less useful for predicting cognitively complex tasks than simpler ones. Consequently, the parameters of the Teamwork KSA Test that are useful in predicting performance of teams working in cognitively complex probably require additional study.

Situated judgment tests have also been used to test more specific elements of team performance. For example, Mumford, Van Iddekinge, Morgeson, and Campion (2008) have recently adopted this approach to assess knowledge of team roles. Their data suggest that the assessment is effective in predicting both team role performance and overall team performance.

16.4.3 Observation

Although not frequently used to assess individual competencies for teamwork, there are some data to suggest that observation of performance might be a useful way to assess the required KSAs. For example, Driskell and Salas (1997) describe an observational approach to infer collective orientation. This approach presents participants with a problem and observes the manner in which they consider information from other teammates. The patterns of information used are the basis of subsequent categorizations about collective orientation. These categorizations were effective in predicting subsequent performance in a laboratory task similar to the categorization task.

Observational techniques are not often studied empirically, but the use of such techniques to assess an individual's competencies for teamwork is sometimes used in practice. For example, in the airline industry, candidates for line duty are usually required to perform a simulated flight scenario specifically tailored to assess various teamwork competencies such as leadership and communication. Failure to demonstrate these teamwork skills is equally as disqualifying as deficiencies in technical skills. This type of approach might be helpful in determining individual competencies for teamwork, especially if combined with an embedded events approach like the TARGETS methodology described by Fowlkes et al. (1994). This technique requires the assessor to identify triggering events for the skills of interest that are appropriate in the context of the scenario. These events are then woven into the narrative of the scenario such that an adequate number of triggers exist for each skill. A companion observation scale is then constructed and an observer records the trainee's success at accomplishing the targeted behaviors.

16.5 Interventions Aimed at Cognitive Readiness for Team Performance

For better or worse, cognitive readiness for teamwork is only a useful construct if it leads to better team performance in the operational environment. We believe that the value of identifying and measuring cognitive readiness for teamwork is that it can be modified prior to team performance. Referring back to Table 16.1, the final column indicates what we believe are the best interventions that can be applied in this regard, including team selection, team composition, and team training. These approaches are detailed in the next section.

16.5.1 Selection

One proposed use of team cognitive readiness evaluations is to use this information to help select personnel to serve in team-based environments (Paris, Salas, & Cannon-Bowers, 2000). An obvious mechanism for individuals to gain competence in teamwork skills is through past experience, so it is conceivable that some individuals will have higher degrees of readiness for team performance than others. Hence, one could identify individuals that *do not* possess the appropriate readiness levels, and assign them to tasks that do not require coordination with other workers. However, as the modern workplace becomes increasingly complex, this approach seems less and less viable. The simple fact is that there are relatively few workplaces that no longer require some degree of coordination. Further, there seems to be a general consensus that team competencies can be trained effectively, leading training to be a preferred application of the data. On the other hand, when time and/or resources are short, it may be viable to identify those individuals in the workforce who already possess the requisite skills.

16.5.2 Team Composition

Perhaps a more practical application of cognitive readiness for teamwork is in the area of team composition. It has been demonstrated that teamwork skills account for more of the variance in team performance than does actual expertise in a member's specific task (Baker & Salas, 1992; Stevens & Campion, 1997). Consequently, one application of these data might be to purposefully compose teams to arrive at a "team mean." The goal of this approach is to identify team members with high ability on one competency to account for members with lower ability on that competency (Hollenbeck, DeRue, & Guzzo, 2004). It may be fruitful, for example, to try different combinations of potential team members until the team mean for each competency exceeds a certain threshold. In so doing, the team should possess an adequate level of competence on each of the critical competencies.

However, while this type of approach is intuitively appealing, data suggest that the method of aggregating the individual team members' scores to arrive at the best solution is complicated. For example, in some cases, one team member holding the competency may be enough to ensure effective teamwork. In other cases, it may be necessary for all team members to be high on a competency (so that adding scores would be best).

In determining which method of aggregation is best, LePine, Hollenbeck, Ilgen, and Hedlund (1997) demonstrated that the appropriate algorithm might vary as a function of the interdependence of the team. Similarly, Barrick, Stewart, Neubert, and Mount (1998) also demonstrated that the best approach to composing teams might be different for each competency considered. Therefore, we conclude that while the use of competency data to compose teams might be beneficial, there is a requirement for more research to guide how best to use these data before the promise of this approach can be realized. In particular, it would be useful to determine not only which components of cognitive readiness for teamwork are necessary before an individual is prepared to enter a team but also how the individual readiness levels of various team members combine to arrive at an overall "team readiness" score.

16.5.3 Training

The most promising intervention for increasing cognitive readiness for team performance is through training. We address two approaches to training in the following sections: training for specific skills that we believe underlie cognitive readiness for team performance and multiple skills training that addresses several components at once. The specific skills training descriptions are broken out according to the skill they address (see Table 16.1 for a summary).

16.5.3.1 Training for Back-up Behaviors

One approach that seems effective in encouraging back-up behaviors is cross-training. *Cross-training* refers to a class of team training interventions that requires team members to experience some or all of their teammates' roles. Blickensderfer, Cannon-Bowers, and Salas (1998) hypothesized that cross-training helps the team develop a "shared mental model," enabling other coordination behaviors. One positive effect of cross-training seems to be improved back-up behavior (Marks, Sabella, Burke, & Zaccaro, 2002). By experiencing the role of their teammate, the participant can better understand what that teammate needs and is more likely to provide it at an appropriate time. In addition, being able to experience the task from another perspective can help team members understand how and when back-up is most essential.

16.5.3.2 Training Communication Skills

Communication training would appear to be a natural training target to improve team performance. However, there are relatively few empirical evaluations of training programs designed to improve communication skills. One exception is Salas, Cannon-Bowers, and Johnston (1997), who described a program for Navy personnel called Team Adaptation and Coordination Training (TACT) that targeted specific communications required in Navy combat information centers. They reported that the program resulted in more effective communication flow in simulated missions.

Yedidia et al. (2003) hypothesized that communication training would aid medical students as they transitioned from individual to team-based work. They found that a group that received communication training performed better not just on communication indices, but on a variety of other performance measures.

Given these results, the notion that communication skills training can be useful is promising. However, it is not clear what level of detail in communication should be the goal of communication training. Despite the identification of apparently important content categories or communication sequences, it is not at all clear that these behaviors can be effectively trained. For example, Sigel and Frederman (1973) observed a large number of helicopter crews in order to extract meaningful communication clusters. Although they were able to identify clusters that reliably discriminated between performance levels, they were unable to successfully train these behaviors. Similarly, Lassiter, Vaughn, Smaltz, Morgan, and Salas (1990) report that although they were able to use training to increase specific communication behaviors, this behavioral change was not associated with improved performance. It is likely that patterns of communication are fairly ingrained and are not likely to change in response to the relatively brief interventions described here. The training of communication skills is an area that is sorely in need of additional study.

16.5.3.3 Interpersonal Skills

A variety of approaches have been used in an attempt to improve interpersonal skills for team members as a means to enhance their subsequent team performance. These programs typically include one or more behavioral clusters including goal setting, interpersonal relations, role clarification, and problem solving (Salas, Rozell, Mullen, & Driskell, 1999). The results of a meta-analysis by Salas, Rozell, et al. (1999) indicated that these interpersonal skills were not reliably associated with team performance. However, Bradley, White, and Mennecke (2003) considered the lifespan of teams in their review and report that interpersonal skills are much more likely to be important for teams with a long lifespan than a shorter one. Hence, it may be viable to consider training for communication skills as a means to enhance cognitive readiness for team performance given that operational teams typically have longer lifespan.

Other researchers have focused on specific interpersonal skills that are thought to be critical for team performance. One such competency is assertiveness. Early studies of critical teamwork mishaps revealed that subordinates were often so intimidated by their leaders that they failed to challenge them, even when the results were disastrous (Foushee, 1984; Prince & Salas, 1993). In response to this problem, researchers developed programs designed to help team members become confident and competent at asserting their viewpoints appropriately. Smith-Jentsch, Salas, and Baker (1996) demonstrated that assertiveness is a trainable competency. Furthermore, Jentsch, Barnett, Bowers, and Salas (1999) demonstrated that skill-based assertiveness training could be further improved by adding a metacognitive training component.

Taken together, the research on training interpersonal skills as a precursor for effective teamwork is encouraging. Clearly, more validation efforts are needed, but, as described previously, at least some results suggest that cognitive readiness for teamwork can benefit from training in interpersonal skills.

16.5.3.4 Improving Teamwork Attitudes

As described previously, the CMAQ can be used to identify specific teamwork attitudes that are inappropriate. Training programs can then be tailored to correct these faulty beliefs. Several studies suggest that training interventions can be effective in changing attitudes towards teamwork (cf. Bowers, Jentsch, & Salas, 2000; O'Connor & Flin, 2003; Salas & Bowers, 2000). However, it is not clear how much these training gains actually translate to improved team performance. Although these data are compelling, it is important to understand that there may be factors that limit the translation of these positive attitudes into behavioral change. For example, Grogan et al. (2004) have recently demonstrated that an 8-h training program was effective in improving the teamwork attitudes of a group of medical employees. However, O'Daniel and Rosenstein (2008) report that after CRM training, compliance with recommended teamwork behaviors was only approximately 60 % in a similar medical sample. It seems that one way to improve these results is to combine attitudes training with training in specific teamwork skills (Salas, Fowlkes, Stout, Milanovich, & Prince, 1999).

16.5.3.5 Leadership

Some researchers have attempted to improve leadership behaviors by targeting the specific behaviors that have been associated with effective team leadership. For example, Tannenbaum, Smith-Jentsch, and Behson (1998) tested an approach to team leader training that emphasized the delivery of effective pre-briefs (before a hands-on exercise) and debriefs (after a hands-on exercise). According to these authors, pre-briefs are essential opportunities for leaders to guide the team in planning, setting mutual goals, developing contingency plans, and clarifying roles and responsibilities. Debriefs also serve multiple purposes; to provide specific feedback on team performance, encourage active team member participation, emphasize teamwork as well as taskwork, and accept feedback from team members.

Moreover, Tannenbaum et al.'s results suggested that this type of training improved not only the behavior of the team leaders but also the overall effectiveness of the teams they lead. Marks, Zaccaro, and Mathieu (2000) demonstrated a similar positive effect of pre-briefings for student teams performing a laboratory task. Hence, it appears that training team leadership skills is a promising mechanism to enhance cognitive readiness for teamwork. Certainly, results bear the notion that leadership skills are an important aspect of team performance and that training can successfully enhance such skills. This conclusion is supported by the recent meta-analysis by Stewart (2006).

16.5.3.6 Multiple Skills Training

Several researchers have developed training programs designed to improve multiple aspects of team performance at once. Salas et al. (Salas, Fowlkes, et al., 1999; Salas, Prince, et al., 1999) have articulated the elements that should be in a broad-based team training program. These include: delivery of the declarative information that team members might not possess, demonstration of critical behaviors, the opportunity to practice newly learned skills in a realistic environment, and feedback about one's performance.

Oftentimes, these training programs are designed to meet the particular needs of a specific type of team. For example, Salas, Prince, et al. (1999) developed a program to satisfy the teamwork needs of naval aviators. They provided specific skill training on several teamwork factors, such as leadership and communication. Evaluations of the program demonstrate skill gains ranging from 8 to 20 % (Salas, Fowlkes, et al., 1999; Salas, Prince, et al., 1999).

Smith-Jentsch, Cannon-Bowers, Tannenbaum, and Salas (2008) developed a multifaceted team training approach that emphasizes guided self-correction. Their approach provides a structured debriefing designed to facilitate shared mental models and to increase the likelihood that team members will discuss critical performance deficiencies and work to correct them. Empirical data indicate that this approach is effective in improving teamwork, communication, and supportive behaviors such as back-up (Smith-Jentsch et al., 2008).

However, it also appears that "generic" team skills training can offer benefits for subsequent team performance. This type of training may be appropriate for those teams that may change tasks frequently (Ellis, Bell, Ployhart, Ilgen, & Hollenbeck, 2005). For example, Pritchard and her colleagues demonstrated that this type of generic team training resulted in higher learning scores for college students that had to work together in a problem-solving task (Pritchard et al., 2006).

Obviously, if training programs can be designed to effectively deliver instruction on multiple cognitive readiness factors for teamwork at once, this is a more efficient strategy than isolating skills and training them separately. Perhaps the most fruitful area to pursue in this regard is specification of skill sets that can be meaningfully clustered so that they can be trained simultaneously. In addition, it may be viable to begin training with individual skills and the target (which are presumably simpler) and then work up to more complex situations where multiple skills can be assessed and trained.

16.6 Conclusion and Future Directions

As the importance of teamwork increases in the modern workplace, so does the need to identify team members who possess the appropriate competencies to perform effectively in these teams. However, as should be clear from the discussion here, our knowledge about the nature of these competencies, their measurement, and how best to use these data is far from complete. There is clearly a need for more research dedicated to this important topic.

That said, a better understanding of individual team competencies might allow us to realize the vision for team training that was described by team training researchers more than a decade ago. For example, Cannon-Bowers and Salas (1998b) described a strategy for team training that emphasized the need for the training development to be based on a thorough training needs analysis. Although training developers have done this, it has largely been at a fairly global level. That is, the training needs of all team members are observed and distilled into a training program that is then delivered to each individual, regardless of whether they possess the competency. While this approach has been largely effective, it would certainly be more efficient to tailor training more specifically to the needs of the trainee. It may be that the measurement of individual teamwork competencies will provide the input for this type of training approach. An example of this approach is provided next.

16.6.1 A Possible Future Application: Tailored Scenario-Based Training

Scenario-based training is essentially a multiple-competency training approach that emphasizes practice of new skills in a simulated task. This type of training is typically deployed with the help of a task simulator, but can be used in lower-fidelity simulations such as a role play exercises. Scenario-based training has been used for the training of technical skills for years, across many domains (e.g., see Cannon-Bowers, Bowers, & Sanchez, 2008). However, it has only been more recently that researchers have indicated that it might also be especially important for the training of teamwork skills as well. The key to scenario-based training for teams is the creation of scenarios that provide an appropriate context for demonstrating the targeted behaviors (Fowlkes et al., 1994; Oser, Cannon-Bowers, Salas, & Dwyer, 1999).

The scenario-based training approach appears to hold tremendous promise for improving team performance. This approach has been used in a variety of settings, including police work (Saus et al., 2006), military settings (Cuevas, Fiore, Bowers, & Salas, 2004), and aviation (Strater & Bolstad, 2008). Scenario-based training also seems to be popular and effective in the training of medical teams (Wallin et al., 2007).

A challenge to implementing the tailored scenario-based training approach is the ability to quickly create valid scenarios based on the competency levels of the participants. Most scenarios are constructed by instructional designers working in concert with subject matter experts. This process can take days or even weeks. However, scientists are developing automated scenario generation systems that might be helpful in meeting this challenge. For example, researchers at the University of Central Florida developed the Rapidly Reconfigurable Lone Oriented Events (RRLOE) generator for the Federal Aviation Administration (Bowers et al., 1997).

RRLOE creates teamwork evaluation scenarios for use in commercial aviation that take far less time and effort to develop than in traditional scenario-based training. Employing user-friendly interfaces and sophisticated algorithms, RRLOE provides viable scenarios and could be easily adapted for use in training (vs. assessment). Other researchers have discussed the potential for using new programming approaches to allow these scenarios to be created with greater attention to individual differences (Schatz, Bowers, & Nicholson, 2009), which might be helpful in creating tailored training.

16.6.2 Conclusions

The notion of cognitive readiness derives its power from the idea that a set of measurable skills can be identified that will predict an individual's likely success in a complex task environment. This provides a mechanism to develop assessments of how prepared an individual is to confront a task and can highlight cases that require intervention. Moreover, cognitive readiness appears to be composed of (at least) some factors that are malleable—that is, they can be improved or changed prior to performance in the environment. This implies that the likelihood of success can be enhanced by targeting cognitive readiness skills prior to deployment.

Applying these notions to the team level is equally powerful. We have suggested that a set of measureable, malleable factors exist that can help us predict whether an individual is ready to be a member of a team. We have also identified likely components of such a construct. Moreover, we have suggested that there are specific measurement approaches and interventions that are appropriate for various cognitive readiness for team performance components. In doing so, we hope that others will continue to develop these ideas so that a body of empirical findings can be amassed.

Acknowledgments The work reported herein was partially supported by a grant from the Office of Naval Research, Award Number N00014-08-1-0126. The findings and opinions expressed here do not necessarily reflect the positions or policies of the Office of Naval Research.

References

Alavi, J., & McCormick, J. (2007). Measurement of vertical and horizontal idiocentrism and allocentrism in small groups. *Small Group Research, 38*(4), 556–564.

Alavi, S. B., & McCormick, J. (2004). A cross-cultural analysis of the effectiveness of the learning organization model in school contexts. *International Journal of Educational Management, 18*(7), 408–416.

Baker, D. P., & Salas, E. (1992). Principles for measuring teamwork skills. *Human Factors, 34*(4), 469–475.

Barnes, C., Hollenbeck, J., Wagner, D., DeRue, D., Nahrgang, J., & Schwind, K. (2008). Harmful help: The costs of backing-up behavior in teams. *Journal of Applied Psychology, 93*(3), 529–539.

Barrick, M. R., Stewart, G. L., Neubert, M. J., & Mount, M. K. (1998). Relating member ability and personality to work-team processes and team effectiveness. *Journal of Applied Psychology, 83*(3), 377–391.

Bass, B. M. (1990). *Bass and Stogdill's handbook of leadership*. New York: Free Press.

Blickensderfer, E., Cannon-Bowers, J. A., & Salas, E. (1998). Cross-training and team performance. In J. A. Cannon-Bowers & E. Salas (Eds.), *Making decisions under stress: Implications for individual and team training* (pp. 299–311). Washington, DC: American Psychological Association.

Bowers, C. A., Braun, C. C., & Morgan, B. B. (1997). Team workload: Its meaning and measurement. In M. T. Brannick, E. Salas, & C. Prince (Eds.), *Team performance assessment and measurement: Theory, methods, and applications* (pp. 85–108). Mahwah, NJ: Erlbaum.

Bowers, C. A., Jentsch, F., & Salas, E. (2000). Establishing aircrew competencies: A comprehensive approach for identifying CRM training needs. In H. F. O'Neil & D. H. Andrews (Eds.), *Aircrew training methods and assessment* (pp. 67–83). Mahwah, NJ: LEA.

Bradley, J., White, B. J., & Mennecke, B. E. (2003). Teams and tasks: A temporal framework for the effects of interpersonal interventions on team performance. *Small Group Research, 34*(3), 353–387.

Brown, S. L., & Eisenhardt, K. M. (1995). Product development: Past research, present findings, and future directions. *Academy of Management Review, 20*, 343–378.

Burke, C., Stagl, K., Klein, C., Goodwin, G., Salas, E., & Halpin, S. (2006). What type of leadership behaviors are functional in teams? A meta-analysis. *The Leadership Quarterly, 17*(3), 288–307.

Burke, R. J., Weir, T., & Duncan, G. (1976). Informal helping relationships in work organizations. *Academy of Management Journal, 19*, 370–377.

Campion, M. A., Medsker, G. J., & Higgs, A. C. (1993). Relations between work group characteristics and effectiveness: Implications for designing effective work groups. *Personnel Psychology, 46*, 823–850.

Cannon-Bowers, J. A., & Bowers, C. (2010). Team development and functioning. In S. Zedeck (Ed.), *APA handbook of industrial and organizational psychology* (Vol. 1, pp. 597–650). Washington, DC: American Psychological Association.

Cannon-Bowers, J. A., Bowers, C. A., & Sanchez, A. (2008a). Using synthetic learning environments to train teams. In V. I. Sessa & M. London (Eds.), *Work group learning: Understanding, improving and assessing how groups learn in organizations* (pp. 315–346). New York, NY: Taylor & Francis Group/Lawrence Erlbaum Associates.

Cannon-Bowers, J. A., & Salas, E. (1998a). Team performance and training in complex environments: Recent findings from applied research. *Current Directions in Psychological Science, 7*(3), 83–87.

Cannon-Bowers, J. A., & Salas, E. E. (1998b). *Making decisions under stress: Implications for individual and team training*. American Psychological Association.

Cannon-Bowers, J. A., Salas, E., Tannenbaum, S., & Mathieu, J. (1995). Toward theoretically based principles of training effectiveness: A model and initial empirical investigation. *Military Psychology, 7*(3), 141–164.

Christensen, C., Laerson, J. R., Jr., Abbott, A., Ardolino, A., Franz, T., & Pfeiffer, C. (2000). Decision making of clinical teams: Communication patterns and diagnostic error. *Medical Decision Making, 20*, 45–50.

Crichton, M. (2005). Attitudes to teamwork, leadership, and stress in oil industry drilling teams. *Safety Science, 43*, 676–695.

Cuevas, H. M., Fiore, S. M., Bowers, C. A., & Salas, E. (2004). Fostering constructive cognitive and metacognitive activity in computer-based complex task training environments. *Computers in Human Behavior, 20*(2), 225–241.

DeChurch, L., & Haas, C. (2008). Examining team planning through an episodic lens: Effects of deliberate, contingency, and reactive planning on team effectiveness. *Small Group Research, 39*(5), 542–568.

Driskell, J. E., & Salas, E. (1997). Collective behavior and team performance. In D. F. Russ-Eft, H. S. Preskill, C. Sleezer, D. F. Russ-Eft, H. S. Preskill, & C. Sleezer (Eds.), *Human resource development review: Research and implications* (pp. 206–222). Thousand Oaks, CA: Sage.

Ellis, A., Bell, B., Ployhart, R., Ilgen, D., & Hollenbeck, J. (2005). An evaluation of generic teamwork skills training with action teams: Effects on cognitive and skill-based outcomes. *Personnel Psychology, 58*(3), 641–672.

Flin, R., Fletcher, G., McGeorge, P., Sutherland, A., & Patey, R. (2003). Anaesthetists' attitudes to teamwork and safety. *Anaesthesia, 58*, 233–242.

Flin, R., O'Connor, P., & Mearns, K. (2002). Crew resource management. Improving teamwork in high reliability industries. *Team Performance Management, 8*, 68–78.

Foushee, H. C. (1984). Dyads and triads at 35,000 feet. Factors affecting group process and aircrew performance. *American Psychologist, 38*, 885–893.

Fowlkes, J., Lane, N., Salas, E., Franz, T., & Oser, R. (1994). Improving the measurement of team performance: The targets methodology. *Military Psychology, 6*(1), 47–61.

Gregorich, S., Helmreich, R., & Wilhelm, J. (1990). The structure of cockpit management attitudes. *Journal of Applied Psychology, 75*, 682–690.

Grogan, E. L., Stiles, R. A., France, D. J., Speroff, T., Morris, J. A., Nixon, B., et al. (2004). The impact of aviation-based teamwork training on the attitudes of health-care professionals. *Journal of the American College of Surgeons, 199*(6), 843–848.

Guzzo, R., & Dickson, M. (1996). Teams in organizations: Recent research on performance and effectiveness. *Annual Review of Psychology, 47*, 307–338.

Hackman, J. R., & Wageman, R. (2005). When and how team leaders matter. In B. Staw & R. Moreland (Eds.), *Research in organization behavior* (Vol. 26). Amsterdam: Elsevier.

Harris, T. E., & Sherblom, J. C. (2002). *Small group and team communication*. New York: Allyn & Bacon.

Helmreich, R. (1984). Cockpit management attitudes. *Human Factors, 26*(5), 583–589.

Helmreich, R., & Merritt, A. C. (1998). *Culture at work in aviation and medicine. National, organizational and professional influences*. Aldershot, UK: Ashgate.

Helmreich, R., Sexton, J. B., & Merritt, A. C. (1997). *The operating room management attitudes questionnaire (ORMAQ)* (University of Texas Aerospace Crew Research Project Technical Report 97-6). Austin, TX: University of Texas.

Helmreich, R. L., & Foushee, H. C. (1993). Why crew resource management? Empirical and theoretical bases of human factors training in aviation. In E. L. Wiener, B. G. Kanki, & R. L. Helmreich (Eds.), *Cockpit resource management*. New York: Academic.

Hertel, G., Konradt, U., & Voss, K. (2006). Competencies for virtual teamwork: Development and validation of a web-based selection tool for members of distributed teams. *European Journal of Work and Organizational Psychology, 15*(4), 477–504.

Hirschfeld, R., Jordan, M., Feild, H., Giles, W., & Armenakis, A. (2006). Becoming team players: Team members' mastery of teamwork knowledge as a predictor of team task proficiency and observed teamwork effectiveness. *Journal of Applied Psychology, 91*(2), 467–474.

Hollenbeck, J. R., DeRue, D., & Guzzo, R. (2004). Bridging the gap between I/O research and HR practice: Improving team composition, team training, and team task design. *Human Resource Management, 43*(4), 353–366.

Jentsch, F., Barnett, J., Bowers, C. A., & Salas, E. (1999). Who is flying this plane anyway? What mishaps tell us about crew member role assignment and air crew situation awareness. *Human Factors, 41*(1), 1–14.

Kozlowski, S. W. J., Gully, S. M., Salas, E., & Cannon-Bowers, J. A. (1996). Team leadership and development: Theory, principles, and guidelines for training leaders and teams. In M. Beyerlein, S. Beyerlein, & D. Johnson (Eds.), *Advances in interdisciplinary studies of work teams* (pp. 253–292). Greenwich, CT: JAI.

Lassiter, D. L., Vaughn, J. S., Smaltz, V. E., Morgan, B. B., & Salas, E. (1990). A comparison of two types of training interventions on team communication performance. In D. Woods & E. Roth (Eds.), *Proceedings of the human factors society 34th annual meeting* (pp. 1372–1376). Santa Monica, CA: Human Factors and Ergonomics Society.

LePine, J. A., Hollenbeck, J. R., Ilgen, D. R., & Hedlund, J. (1997). Effects of individual differences on the performance of hierarchical decision-making teams: Much more than g. *Journal of Applied Psychology, 82*, 803–811.

Lingard, L., Espin, S., Whyte, S., Regehr, G., Baker, G., Reznick, R., et al. (2004). Communication failures in the operating room: An observational classification of recurrent types and effects. *Quality & Safety in Health Care, 13*(5), 330–334.

MacMillan, J., Entin, E., & Serfaty, D. (2004). Communication overhead: The hidden cost of team cognition. In E. Salas & S. M. Fiore (Eds.), *Team cognition: Understanding the factors that drive process and performance* (pp. 61–82). Washington, DC: American Psychological Association.

Marks, M. A., Mathieu, J. E., & Zaccaro, S. J. (2001). A temporally based framework and taxonomy of team processes. *Academy of Management Review, 26*, 356–376.

Marks, M. A., Sabella, M. J., Burke, C., & Zaccaro, S. J. (2002). The impact of cross-training on team effectiveness. *Journal of Applied Psychology, 87*(1), 3–13.

Marks, M., Zaccaro, S., & Mathieu, J. (2000). Performance implications of leader briefings and team-interaction training for team adaptation to novel environments. *Journal of Applied Psychology, 85*(6), 971–986.
Mathieu, J., Maynard, M. T., Rapp, T., & Gilson, L. (2008). Team effectiveness 1997–2007: A review of recent advancements and a glimpse into the future. *Journal of Management, 34*, 410–476.
Mathieu, J., & Shulze, W. (2006). The influence of team knowledge and formal plans on episodic team process-performance relationships. *Academy of Management Journal, 49*, 605–619.
Maynard, M. T., Mathieu, J., Marsh, W., & Ruddy, T. (2007). A multilevel investigation of the influences of employees' resistance to empowerment. *Human Performance, 20*, 147–171.
McClough, A. C., & Rogelberg, S. G. (2003). Selection in teams: An exploration of the teamwork knowledge, skills, and ability test. *International Journal of Selection and Assessment, 11*(1), 56–66.
Mesmer-Magnus, J., & DeChurch, L. (2009). Information sharing and team performance: A meta-analysis. *Journal of Applied Psychology, 94*, 535–546.
Miller, D. L. (2001). Reexamining teamwork KSAs and team performance. *Small Group Research, 32*(6), 745–766.
Mohrman, S. A., & Cohen, S. G. (1995). When people get out of the box: New relationships, new systems. In A. Howard (Ed.), *The changing nature of work* (pp. 365–410). San Francisco: Jossey-Bass.
Morey, J. C., Simon, R., Jay, G. D., Wears, R. L., Salisbury, M., Dukes, K. A., et al. (2002). Error reduction and performance improvement in the emergency department through formal teamwork training: Evaluation results of the MedTeams project. *Health Services Research, 37*(6), 1553–1581.
Morgan, B., Salas, E., & Glickman, A. (1993). An analysis of team evolution and maturation. *The Journal of General Psychology, 120*(3), 277–291.
Morgeson, F., Reider, M., & Campion, M. (2005). Selecting individuals in team settings: The importance of social skills, personality characteristics, and teamwork knowledge. *Personnel Psychology, 58*(3), 583–611.
Morrison, J., & Fletcher, J. D. (2002). *Cognitive readiness* (IDA Paper P-3735). Alexandria, VA: Institute for Defense Analyses.
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. *The Journal of Applied Psychology, 93*(2), 250–267.
O'Connor, P., & Flin, R. (2003). Crew resource management training for offshore oil production teams. *Safety Science, 41*, 591–609.
O'Daniel, M., & Rosenstein, A. (2008). Professional communication and team coordination. In R. Hughes (Ed.), *Patient safety and quality*. Washington, DC: AHRQ.
Oser, R. L., Cannon-Bowers, J. A., Salas, E., & Dwyer, D. J. (1999). Enhancing human performance in technology-rich environments: Guidelines for scenario-based training. In E. Salas (Ed.), *Human/technology interaction in complex systems* (pp. 175–202). Stamford, CT: JAI Press.
Paris, C., Salas, E., & Cannon-Bowers, J. (2000). Teamwork in multi-person systems: A review and analysis. *Ergonomics, 43*(8), 1052–1075.
Patrashkova-Volzdoska, R., McComb, S., Green, S., & Compton, W. (2003). Examining a curvilinear relationship between communication frequency and team performance in cross-functional project teams. *IEEE Transactions on Engineering Management, 50*(3), 262–269.
Ployhart, R., & Bliese, P. (2006). I-ADAPT theory: Conceptualizing the antecedents, consequences, and measurement of individual differences in adaptability. In E. Salas (Ed.), *Advances in human performance and cognitive engineering research*. Bingley, UK: Emerald.
Porter, C. (2005). Goal orientation: Effects on backing up behavior, performance, efficacy, and commitment in teams. *Journal of Applied Psychology, 90*(4), 811–818.
Porter, C., Hollenbeck, J., Ilgen, D., Ellis, A., West, B., & Moon, H. (2003). Backing up behaviors in teams: The role of personality and legitimacy of need. *Journal of Applied Psychology, 88*(3), 391–403.

Prichard, J. S., Bizo, L. A., & Stratford, R. J. (2006). The educational impact of team-skills training: Preparing students to work in groups. *British Journal of Educational Psychology, 76*(1), 119–140.
Prince, C., & Salas, E. (1993). Training and research for teamwork in the military aircrew. In E. L. Wiener, B. G. Kanki, & R. L. Helmreich (Eds.), *Cockpit resource management* (pp. 337–366). Orlando, FL: Academic.
Pulakos, E., Arad, S., Donovan, M., & Plamondon, K. (2000). Adaptability in the workplace: Development of a taxonomy of adaptive performance. *Journal of Applied Psychology, 85*(4), 612–624.
Pulakos, E., Dorsey, D., & White, S. (2006). Adaptability in the workplace: Selecting an adaptive workforce. In C. S. Burke, L. G. Pierce, & E. Salas (Eds.), *Understanding adaptability: A prerequisite for effective performance within complex environments* (pp. 41–71). Amsterdam, Netherlands: Elsevier.
Salas, E., & Bowers, C. A. (2000). The design and delivery of crew resource management training: Exploiting available resources. *Human Factors, 42*(3), 490–511.
Salas, E., Cannon-Bowers, J. A., & Johnston, J. (1997). How can you turn a team of experts into an expert team?: Emerging training strategies. In C. E. Zsambok, G. Klein, C. E. Zsambok, & G. Klein (Eds.), *Naturalistic decision making* (pp. 359–370). Hillsdale, NJ: Erlbaum.
Salas, E., Fowlkes, J., Stout, R., Milanovich, D., & Prince, C. (1999). Does CRM training improve teamwork skills in the cockpit? Two evaluation studies. *Human Factors, 41*(2), 326–343.
Salas, E., Prince, C., Bowers, C., Stout, R., Oser, R., & Cannon-Bowers, J. (1999). A methodology for enhancing crew resource management training. *Human Factors, 41*(1), 161–172.
Salas, E., Rozell, D., Mullen, B., & Driskell, J. E. (1999). The effect of team building on performance: An integration. *Small Group Research, 30*(3), 309–329.
Saus, E.-R., Johnsen, B. H., Eid, J., Riisem, P. K., Andersen, R., & Thayer, J. F. (2006). The effect of brief situational awareness training in a police shooting simulator: An experimental study. *Military Psychology, 18*(Suppl), S3–S21.
Schatz, S., Bowers, C., & Nicholson, D. (2009). Advanced situated tutors: Design, philosophy, and a review of existing systems. In *Proceedings of the 53rd Annual Meeting of the Human Factors and Ergonomics Society.* Santa Monica, CA: HFES.
Smith-Jentsch, K. A., Cannon-Bowers, J. A., Tannenbaum, S. I., & Salas, E. (2008). Guided team self-correction: Impacts on team mental models, processes, and effectiveness. *Small Group Research, 39*(3), 303–327.
Smith-Jentsch, K. A., Salas, E., & Baker, D. P. (1996). Training team performance-related assertiveness. *Personnel Psychology, 49*(4), 909–936.
Stajkovic, A. D., Lee, D., & Nyberg, A. J. (2009). Collective efficacy, group potency, and group performance: Meta-analyses of their relationships, and test of a mediation model. *Journal of Applied Psychology, 94*(3), 814–828.
Stevens, M. J., & Campion, M. A. (1997). *Teamwork-KSA test: Examiners manual*. Rosemont, IL: McGraw-Hill/London House.
Stevens, M. J., & Campion, M. A. (1999). Staffing work teams: Development and validation of a selection test for teamwork settings. *Journal of Management, 25*(2), 207–228.
Stewart, G. (2006). A meta-analytic review of the relationship among team design features and team performance. *Journal of Management, 32*, 29–55.
Stone, R., & Bailey, J. (2007). Team conflict self-efficacy and outcome expectancy of business students. *The Journal of Education for Business, 82*, 56–62.
Stout, R. J., Driskell, J. E., & Salas, E. (1997). Collective orientation and team performance. In *Proceedings of the 41st Annual Meeting of the Human Factors and Ergonomics Society* (pp. 1190–1194). Santa Monica, CA: HFES.
Strater, L., & Bolstad, C. (2008). Simulation-based situation awareness training. In D. Vincenzi (Ed.), *Human factors in simulation and training*. Boca Raton, FL: CRC Press.
Svensson, J., & Andersson, J. (2006). Speech acts, communication problems, and fighter pilot team performance. *Ergonomics, 49*(12/13), 1226–1237.

Tannenbaum, S. I., Smith-Jentsch, K. A., & Behson, S. J. (1998). Training team leaders to facilitate team learning and performance. In J. A. Cannon-Bowers & E. Salas (Eds.), *Making decisions under stress: Implications for individual and team training* (pp. 247–270). Washington, DC: American Psychological Association Press.

Urban, J., Bowers, C., Monday, S., & Morgan, B. (1995). Workload, team structure, and communication in team performance. *Military Psychology, 7*(2), 123–139.

Wallin, C.-J., Meurling, L., Hedman, L., Hedegård, J., & Felländer-Tsai, L. (2007). Target-focused medical emergency team training using a human patient simulator: Effects on behaviour and attitude. *Medical Education, 41*(2), 173–180.

Wears, R. L., Perry, S. J., Shapiro, M., Beach, C., Croskerry, P., & Behara, R. (2003). Shift changes among emergency physicians: Best of times, worst of times. In *Proceedings of the 47th Annual Meeting of the Human Factors and Ergonomics Society* (pp. 1420–1423). Santa Monica, CA: HFES.

Weaver, S., Rosen, M., DiazGranados, D., Lazarra, E., Lyons, R., Salas, E., et al. (2010). Does teamwork improve performance in the operating room? A multilevel evaluation. *Joint Commission Journal on Quality and Patient Safety, 36*, 133–142.

Yedidia, M. J., Gillespie, C. C., Kachur, E., Schwartz, M. D., Ockene, J., Chepaitis, A. E., et al. (2003). Effect of communications training on medical student performance. *Journal of the American Medical Association, 290*(9), 1157–1165.

Chapter 17
Impact of Individual Game-Based Training on Team Cognitive Readiness

Talib S. Hussain, Clint Bowers, and Holly Blasko-Drabik

17.1 Introduction

Training systems based on modern computer-game engines have been used to address a number of training needs within the U.S. military and other agencies with varying degree of success (O'Neil & Perez, 2008; Tobias & Fletcher, 2011). A key challenge when designing a training game is to ensure that it will effectively improve the capabilities of the target population—on their real tasks as performed in the real world. A key factor distinguishing how tasks are performed is whether they are done as individuals or as teams on group tasks. Group tasks are tasks in which no one individual can solve the task, and thus cooperation of team members is needed to solve the task. The military in particular has a strong need to ensure that all their warfighters can perform effectively in teams. However, what are effective approaches to addressing this training need? Games have been developed that successfully teach individual skills to individuals, such as operationally relevant language and culture skills (Johnson, 2009; Johnson, Wang, & Wu, 2007) and shipboard damage control skills (Hussain et al., 2010, 2012), and others that successfully teach team-based skills to teams within a multiplayer gaming environment, such as team convoy operations (Diller, Roberts, Blankenship, & Nielsen, 2004; Roberts, Diller, & Schmitt, 2006) and capture the flag activities (Hussain et al., 2008). However, there has been no clear evidence supporting the use of games to improve the performance of teams through individual training.

T.S. Hussain (✉)
Raytheon BBN Technologies, 10 Moulton Street, Cambridge, MA 02138, USA
e-mail: thussain@bbn.com

C. Bowers • H. Blasko-Drabik
Department of Psychology, University of Central Florida,
4000 Central Florida Blvd., Orlando, FL 32816-1390, USA
e-mail: clint.bowers@ucf.edu; holly.blasko@gmail.com

H.F. O'Neil et al. (eds.), *Teaching and Measuring Cognitive Readiness*,
DOI 10.1007/978-1-4614-7579-8_17, © Springer Science+Business Media New York 2014

Cognitive readiness has been defined as the mental preparation (including skills, knowledge, abilities, motivations, and personal dispositions) an individual needs to establish and sustain competent performance in the complex and unpredictable environment of modern military operations (Morrison & Fletcher, 2002). Cognitive readiness skills of individuals and general teaming skills share dependencies (see Bowers & Cannon-Bowers, this volume). The cognitive ability to maintain situational awareness is an important part of team monitoring skills. The cognitive ability of decision-making is an important part of team leadership. The ability of individuals to communicate effectively is an important part of a team's ability to coordinate and build shared mental models. We investigated cognitive readiness skills on individual and team-based damage control tasks at a military recruit training center.

The U.S. Navy Recruit Training Command (RTC) trains over 35,000 recruits per year (Miller, 2012). During an intensive 8-week program, recruits are trained on a number of basic seamanship skills, as well as how to work together effectively in teams. In a capstone event, recruits demonstrate their newly learned capabilities through a number of team-based exercises aboard a physical ship simulation. An analysis of the capstone event concluded that the recruits did not have "sufficient opportunity to adequately practice tasks/behaviors as necessary prior to performance evaluation" (HPC, 2008, p. 3). Based on the resulting recommendations, a training game was developed to improve the recruit's capabilities in the areas of damage control. We hypothesized that a significant improvement in the team performance in the capstone event could be attained by providing training to individuals that would improve their cognitive readiness for shipboard operations. The training game, the Virtual Environments for Ship and Shore Experiential Learning (VESSEL) Damage Control Trainer (DCT), was designed to emphasize the cognitive skills of situational awareness, decision-making, and communication and to help recruits build an effective mental model of shipboard environments and shipboard activities (Hussain et al., 2010, 2012).

To validate our hypothesis, we conducted two studies. In the first study (Bowers, Hussain, Roberts, Cannon-Bowers, & Blair, forthcoming), we evaluated the impact that playing the game had on the performance of individuals in a shipboard damage control transfer task. Significant improvements in individual cognitive readiness and task performance were observed. In the second study, we evaluated the impact that playing the game had on the performance of teams in several damage control and non-damage control shipboard transfer tasks. Significant improvements in team performance were observed, and several indicators of enhanced team cognitive readiness were noted.

We present the design of the game and its emphasis on cognitive skills, introduce each study, and discuss the results of the team study. We discuss a number of lessons that were learned about assessing the performance of teams within a complex real-world environment and provide a set of recommendations for future research on training and assessing team performance and cognitive readiness.

17.2 Background

17.2.1 Navy Recruit Training

The U.S. Navy RTC at Great Lakes, Illinois, the Navy's only boot camp, trains over 35,000 recruits per year (Miller, 2012). Recruits undergo 8 weeks of training, with a heavy emphasis on classroom lectures and drill instruction and a few hands-on exercises (HPC, 2008). At the end of their training, recruits are evaluated in an intense 10-h capstone evaluation event, Battle Stations 21 (BS21). BS21 is based on a physical simulation of an Arleigh Burke class destroyer and includes a simulated dock, ship exterior, and several internal decks. During the event, recruits complete 17 different training scenarios operating in teams of 8–11 in size. The teams are formed at the beginning of the event and maintained throughout the evening. Scenarios include typical shipboard and dockside operations, such as moving stores (i.e., moving boxes of goods or equipment to or from storage compartments), patrolling the ship, securing the ship to the dock or maintaining a watch on the bridge, as well as atypical damage control situations that occur after a simulated explosion while "at sea," such as flooding control, fire fighting, or dealing with injured shipmates. Each team is provided with an experienced facilitator that guides them through all the scenarios during the event. At the beginning of each scenario, the facilitator gives the recruits a verbal briefing of their objectives for that scenario. As the recruits execute the scenario, the facilitator observes and evaluates their performance against a fixed set of criteria for that scenario using a handheld device (HPC, 2008). Although the teams remain constant, during each exercise a different team leader is selected from the team members by the facilitator. The facilitators intervene when safety is an issue but otherwise are instructed to provide minimal guidance during the scenario. Battle Stations 21 represents a high-performance environment, with rapidly evolving scenarios, high time pressure, command and peer pressure, fatigue, and a potentially changing team dynamic as leaders change.

Within BS21, as analyzed in 2008, recruits faced challenges due to limited opportunities to practice the skills they had been taught and the need to learn significant new material once they were in the BS21 exercise; "Recruits do not have the knowledge or ability to accurately complete the… (scenario) criteria as presented without assistance, facilitation, and on the first attempt" (HPC, 2008, p. 14). In particular, these issues could be seen in damage control situations, where a number of critical errors were made due to poor decisions which, in real-life situations, would endanger the ship and its sailors.

17.2.2 Game-Based Training

A particular challenge for training individuals and teams to perform in complex environments is that traditional lecture-based approaches do not lend themselves to the explanation of targeted skills. Rather, novice learners often exit such training with only a superficial understanding of the concepts taught and the relationships between these constructs (Glaser, 1989). Scientists attempting to improve on this state of affairs have argued that experiential learning might be far more effective in creating positive training outcomes. This type of learning requires trainees to acquire new knowledge and skills in meaningful contexts in order to build more accurate mental models of the complex environment (Cannon-Bowers & Bowers, 2009; Kolb, 1984).

Once these models have been created, trainees have a much easier time transferring what they already have learned to new situations, contexts, or problems. For example, in an attempt to increase understanding and compliance with aviation safety rules, a study was conducted wherein pilots were asked to play a scenario in a computer flight simulator game (Molesworth, Bennett, & Kehoe, 2011). After playing one group of pilots completed a relapse-prevention questionnaire (a basic knowledge test) and the second group completed a self-explanation questionnaire (pilots were required to explain their actions and how and why they erred) and the control group only played the game. A week later the pilots returned and played both the original and a new flight scenario. Results showed that the pilots in the self-explanation group showed no safety errors for the repeated scenario and the best performance on the novel one of all three groups.

Despite the advantages of experiential learning, it is often difficult to implement. The use of actual transfer environments is often prohibited by cost. Most organizations do not have the luxury of using their workplaces for training purposes due to the associated loss of productivity. Furthermore, many of these complex environments are often too dangerous to be used by untrained personnel. The risk of harm often far exceeds the benefits of such training.

Training professionals have attempted to respond to these challenges by developing simulations designed to support experiential training. These simulations are often high-fidelity mock-ups that closely resemble the workplace environment. While these simulations are often useful in improving training, they are often subject to the same cost concerns as the actual workplace (Smith & Roehrs, 2009). High-fidelity simulations are not only expensive to acquire, but they often require dedicated training and safety personnel and are expensive to maintain. These expenses are often far higher than is affordable.

Recently, scientists have explored the lower limits of simulator fidelity that still possess training value (Brydges, Carnahan, Rose, Rose, & Dubrowski, 2010; De Giovanni, Roberts, & Norman, 2009). Results of these studies have suggested that substantial training value can be obtained using training approaches delivered on personal computers. Many of these training programs have been developed using computer-game programming tools or even modifying the games themselves (Fletcher, 2009; Jentsch & Bowers, 1998; Scalese, Obeso, & Issenberg, 2007). It has been argued that the effectiveness of these approaches is likely based on the degree to which they are developed in line with recommendations from the science of learning (Cannon-Bowers & Bowers, 2009). One example of technology based on the science of learning is the use of games for training, which has been increasing over the past few years (O'Neil & Perez, 2008; Tobias & Fletcher, 2011). Computer-game-based training systems have a great potential to affect the learning and cognitive readiness of individuals (Koenig, Lee, Iseli, & Wainess, 2009). For instance:

- Games provide interactive experiences in a task-based environment with repeated exposure to important cue patterns. These are important elements in the development of expertise (Chi, Glaser, & Farr, 1988; Bransford, Brown, & Cocking, 1999; Rupp, Gushta, Mislevy, & Shaffer, 2010).
- Games support model-based reasoning by providing a world in which students may manipulate variables, draw and test hypotheses, and compare their mental models with representations in the game world (Cartier & Stewart, 2000; Gentner, 1983; Mayrath, Clarke-Midura, Robinson, & Schraw, 2011; Raghavan, Satoris, & Glaser, 1997; Stewart, Cartier, & Passmore, 2005; Zimmerman, Raghavan, & Sartoris, 2003).
- Games provide opportunities for individuals to experience the consequences of their decisions and improve their ability to identify the cues that support effective decision-making (Swarz et al., 2010).
- Games challenge the student to maintain a heightened awareness of the game environment to identify elements that support or conflict with their tasks in a timely manner (Mayrath et al., 2011).

17.2.3 Team Cognitive Readiness

As discussed in an earlier chapter (see Bowers & Cannon-Bowers, this volume), several cognitive readiness competencies for teamwork may be identified, including situational awareness, communication, and decision-making. Of these, communication in particular has a broad-based impact on team performance (Cannon-Bowers & Bowers, 2011; Christensen et al., 2000; Wears et al., 2003). Communicating with sufficient, but not excessive, frequency and detail is likely to be a key contributor to good team performance (Kennedy & McComb, 2010; MacMillan, Entin, & Serfaty, 2004; Patrashkova-Volzdoska, McComb, Green, & Compton, 2003). As such, training of communication skills should be an effective method of improving team performance.

However, while this is a promising premise (Yedidia et al., 2003), there is limited empirical evidence regarding the positive impact of specific communications training on team performance (Salas, Cannon-Bowers, & Johnston, 1997) and some evidence indicating that training specific communication behaviors to team members may not produce an improvement in team performance (Lassiter, Vaughn, Smaltz, & Morgan, 1990). Thus, there is a need for further research to investigate the impact of communications training on team performance.

Further, while training on a single teamwork skill, such as communication, may provide a benefit to team performance, there is also evidence (Salas, Fowlkes, Stout, Milanovich, & Prince, 1999) that training designed to address multiple cognitive readiness factors for teamwork at once may be particularly effective for improving team performance.

17.2.4 Assessing Team Performance Through Observation

Assessing the cognitive readiness skills related to team performance may be done in several ways, including subjective tests, situated judgment tests, and observation-based protocols (see Bowers & Cannon-Bowers, this volume). Due to the particular constraints of our customer regarding the availability of the recruit subject pool, an observational protocol was required to both minimize disruption of the recruits and maximize the operational validity of the results. In particular, our customer wanted to know the specific impact of playing the game upon the recruit's actual experience in the BS21 event. BS21 scenarios offered repeatable situations within which to observe team performance since all teams experienced the same tasks and constraints. Due to the large size of the BS21 environment, for most scenarios there was ample room for observation by an observer following the team in person. For scenarios involving small compartments, a single person could still observe from within the compartment but needed to coordinate his or her position with the facilitator throughout the scenario to ensure that they were out of the way of the recruits and the facilitator as well as to ensure their own safety (e.g., entered the compartment after the recruits but before the facilitator, stood out of the way of the moving recruits within the compartment and without obstructing the field of view of the facilitator). However, BS21 also provided many challenges for an observational study protocol since the dynamic of the teams changed as the leaders changed, different facilitators provided slightly different guidance to their teams, different teams encountered the scenarios in different orders, and the pace of activities in many scenarios was very fast. In order to minimize the impact of these challenges, we developed our observation protocol with inputs from the trainers at RTC to account for known scenario-specific behaviors and errors, solicited inputs from the facilitators after each scenario was completed in the team study to capture any additional observations, and piloted our protocol.

17.3 The VESSEL Damage Control Trainer Game

The VESSEL DCT is a single-player, 3D, first-person perspective game that provides training on flooding control skills using guided discovery. For example, it provides a narrative to ground the experiences of the student and develops that narrative as the levels progress. The game, based on the Delta3D game engine (Murphy, Rodgers, Guthrie, & McDowell, 2008), takes place inside a simulated Arleigh Burke class destroyer in which the student is able to navigate passageways, enter

Fig. 17.1 Interactive virtual ship environment for damage control training

compartments, interact with objects in the environment, and use repair equipment (e.g., see use of jubilee patch equipment in the middle of Fig. 17.1). In the game, the student performs activities similar to those on a real ship (using the communication system, navigating through the ship, using items).

The game interface is designed to maximize the student's interaction with the environment while providing mechanisms to support instructional guidance and feedback. Each level of the game is a "mission" in which the student is given some initial objectives to achieve. For example, in one of the missions, the player is asked to dress out properly for flooding and investigate a compartment with a suspected problem. Further interactions while playing the level refine and add to those objectives. For example, after discovering a leaking fire main in the compartment, the player is then given the objective to obtain the correct repair equipment and repair the leak. During the mission, the student is given guidance depending upon performance. For instance, after experiencing some difficulty with patching the pipe, the player is provided a brief reminder of the correct procedure for applying a jubilee patch to a leaking pipe (e.g., see the text box at the top-left corner of Fig. 17.1). In a mission debrief at the end, the student is given a detailed performance assessment and allowed to proceed to the next game mission if sufficient skill has been demonstrated. An "individual development plan" in the debrief provides specifics on the

Fig. 17.2 Individual development plan showing summary of student performance

situational awareness, communication, and decision-making aspects of the student's performance (e.g., see Fig. 17.2).

The DCT used in the studies included three levels:

1. A tutorial level that introduces game mechanics and the ship environment.
2. A ship navigation level that teaches how to identify locations in a ship and move from location to location within a ship.
3. A flooding control level that teaches how to investigate for potential damage and repair a leaking fire main and that reinforces making effective observations, reports, and decisions in a damage control situation.

The training provided in DCT is based on what the recruits are taught in the RTC curriculum. The goal of the student is to practice their damage control skills and improve over time. To this end, significant feedback and opportunities for replay are given. Within the game, the student is given specific objectives in each level that must be followed using the appropriate Navy protocols. The ultimate objective within the narrative context of the game is to support the ship in its mission (as it heads to its duty station in the Persian Gulf)—just as a sailor would in real life. The student proceeds through situations that unfold over time. For instance, in the first two levels, the student secures different areas of the ship in preparation for an underway replenishment from another ship. Following those two levels, the student observes as a collision occurs between the two ships during the underway replenishment. In the subsequent flooding control level, the student performs damage control activities in response to that collision—in particular, repairing a fire main that sprung a leak as a result.

The player is given different specific roles across the missions and interacts with disembodied virtual characters via text-based interaction (e.g., Damage Control Central on the other end of the phone) to receive orders, communicate the situation and receive changes in orders. Through these mechanisms, the individual is provided with training that emphasizes cognitive skills in the context of damage control and shipboard operations without explicit training on performance in teams.

Positive and negative feedback and rewards are given in the game as appropriate. As the student makes good choices and achieves objectives, positive feedback statements may pop up and a green bar indicating mission progress may increase. Occasionally a correct dialog response will result in a supportive statement from the other person in the dialog. As the student makes mistakes in a mission, demerits are issued. A demerit provides a pop-up message explaining the mistake and increases a red performance penalty bar by a predetermined amount to indicate the severity of the error. For instance, opening a watertight door without permission may result in a penalty since the player potentially put the ship at risk. Penalties on repeated errors are larger and the associated guidance messages are usually more detailed. It is possible for the student to fail the mission if too many errors are accumulated (i.e., the red bar fills up). Certain key actions in the game (e.g., those that endanger life or ship safety) can also result in an immediate failure—the student is told they have made a "catastrophic failure" and the consequences of those actions may be highlighted using images or video. For instance, shutting a valve on a fire main without permission may lead to the significant injury of another shipmate elsewhere in the ship fighting a fire, as shutting the fire main may result in no water being available to fight a fire elsewhere in the ship (see Fig. 17.3). If students fail a mission, they must replay that mission again from the beginning; they cannot proceed to the next mission in the game until they succeed.

Throughout the game, guidance, feedback, and consequences are used to reinforce the cognitive skills of situational awareness, communication, and decision-making and to assist the student in building appropriate mental models of shipboard operations and damage control priorities. In particular, the game reinforces a basic "Assess-Report-Act" approach to damage control and shipboard operations. Three high-level categories of learning objectives related to cognitive readiness are used as an organizing principle:

- Situational awareness: The ability to recognize cues, interpret cues, and predict consequences
- Communication: The ability to know whom to contact, when to contact, and how to report
- Decision-making: The ability to follow appropriate protocols, follow orders, and take initiative to complete a mission

Fig. 17.3 Catastrophic failure video emphasizing consequences of student's critical error

For situational awareness, the student is encouraged to pay attention to their surroundings, provide complete, accurate descriptions of the initial situation and changes in the situation, and cautioned to consider the consequences of their actions. For communication, the student is taught to pay attention to the need to make reports, to do so in a timely manner, to use standard Navy communication protocols, and to ensure that reports are appropriate, accurate, and complete. For decision-making, the student is reinforced on the correct damage control methods and encouraged to make timely decisions that maximize mission success and minimize danger to themselves and their shipmates. The importance of these skills to the overall construct of cognitive readiness is described by Bowers and Cannon-Bowers (this volume).

Student performance is evaluated based on the dialog choices made by the student, physical choices of the student in the game, whether the student accomplishes certain activities, and how long the student takes to perform certain tasks. At the end of the level, the student is rated on their performance. During this debrief, all of the learning objectives for the mission are grouped by their relevant category (e.g., in Fig. 17.2, the "Assess Flooding Situation" learning objective is in the Situational Awareness category). The student's performance against each learning objective is evaluated using a simple 3-level "stoplight" metric. If no demerits were received for that tasks related to that objective, the student receives a green result. For most objectives, if a single demerit was received on related tasks, a yellow result is given (e.g., see "Follow Safety Protocols" in Fig. 17.2), and if two or more demerits were received on related tasks, a red result is given (e.g., see "Recognize Shipboard Navigation Cues" in Fig. 17.2). For certain objectives that were deemed particularly critical as part of the game design, there is no yellow threshold (i.e., a single demerit results in a red result).

In addition to these per-objective stoplight evaluations, the student also receives a summary ranking from zero to five on their performance (see bottom right of Fig. 17.2). Mission failure (i.e., due to excessive demerits or a catastrophic failure) always receives a rank of zero. In the case of mission success, the rank is determined as follows:

- 1 point is given for completing the mission. This 1 point is guaranteed, regardless of the level of demerits received or time taken.
- 0 to 3 points are assigned based on the ratio of demerits received by the player on the mission compared to the maximum allowable demerits (3 points for no demerits, 2 points for fewer than 30 % of the maximum, 1 point for 30–60 % of maximum, and 0 points for more than 60 % of maximum).
- A point may be added or deducted based on the time taken by the player to complete the mission compared to a base "average" or "poor" time for that mission (+1 point if the player's time was less than the base average time, −1 point if the time was greater than the base poor time, 0 additional points otherwise). The average and poor times were determined for each mission based on playtests of that mission during development. One thing to note is that the playing time is only tracked when the player can move around. The time does not increment when the player is in a dialog window or other information-based window so as not to penalize slower readers.

Thus, for example, a rank of 5 indicates that they completed the mission quickly with no errors, while a rank of 1 indicates that they completed the mission slowly with a lot of errors.

17.4 Transfer Study for Individuals

A validation study (Bowers et al., forthcoming) was conducted to test the impact of the DCT on individual performance within a flooding control test scenario in the Battle Stations 21 environment. Participants consisted of 31 Navy recruits from RTC. All recruits had completed RTC training but had not yet done the BS21 capstone evaluation. Participants were randomly assigned to game (15 participants) or no-game (16 participants) conditions. Recruits in the game condition played the game for 1 h and then performed in a transfer scenario 48 h later. During the 48 h between playing the game and the scenario, recruits carried out their duties as normal. The no-game participants had no extra training and performed in the same transfer scenario.

Table 17.1 Highest performance difference on cognitive readiness skills of game group, compared to no-game group, in the individual and team studies

	Highest performance difference of game group compared to no-game group	
Cognitive readiness skill	Individual study (%)	Team study (%)
Situational awareness	67	41
Communication	80	26
Decision-making	38	43

The transfer scenario was conducted in the same physical ship simulation used in Battle Stations 21 but was different from any formal BS21 scenario. Instead of participating within a team, each subject participated as an individual performing a flooding control scenario similar to the one in the game—such as dressing out in appropriate protective equipment, locating a room with a potential flood, reporting the flood, and repairing the leak. The recruit needed to perform all their tasks with no help from the facilitator or DCC, even if help was solicited. If the subjects placed themselves in potential danger, the facilitator would intervene to redirect the subject accordingly.

The participants were evaluated, using a direct observational protocol, on their individual performance on a number of behaviors related to communications, decision-making, and situational awareness within the ship (see Appendix 1 for observational form used). Communication skills was assessed by recording actions taken and errors made in standard communication protocols that had been trained earlier. Situational awareness was inferred from reactions to cues in the training scenario. Likewise, decision-making was inferred from behaviors in response to scenario demands. The results show a dramatic, significant improvement in the performance of the game group on almost all measures and across all cognitive readiness skills observed (see Table 17.1). This validated that the individual training had a significant impact on individual performance and individual cognitive readiness. In addition to these specific objective measures, the behaviors of the two groups were visibly quite different. The individuals in the control group generally appeared confused as to what they should do and made frequent requests for help. The individuals in the treatment group were generally confident in their actions, made few requests for help, and appeared to be enjoying the challenge of the test.

17.5 Transfer Study for Teams

17.5.1 Methods

A second validation study was conducted to test the impact of the DCT on team performance within BS21 and to validate our hypothesis that individual training that produces improved individual readiness leads to improved team performance.

Participants consisted of 322 Navy recruits from RTC. All recruits were in the final week of their basic training. Recruits had been randomly assigned to divisions prior to the study. Divisions were then randomly assigned to game or no-game conditions. Groups that were in the game condition played the game for approximately 1 h as their schedules permitted in the 2 days before BS21 took place. All groups were then observed in the transfer environment (BS21).

The key performance variables were obtained from the recruits' performance in the usual Battle Stations 21 capstone event. In the study, the recruits' performance in three scenarios was observed to test transfer to damage control and non-damage control situations. In all three scenarios, researchers paid particular attention to specific behavioral indices of competence suggested by subject matter experts. These focused on the core skill areas of communication, situational awareness, and decision-making. For each concept, a few key behavioral indicators were chosen:

- Flooding magazine compartment scenario—The flooding task is the busiest and most complex task researchers observed. The recruits must equip special damage control gear, check and obtain permission to open several doors, properly fix a badly leaking pipe, maintain watertight security of the ship, move nearly 100 rounds of ammunition that weigh around 25 lbs each from the flooding compartment to a neighboring dry compartment, and follow correct damage control protocols. Researchers followed the team into the flooding compartment to observe as they tried to repair the leak and move the ammunition to safety. Researchers captured indicators of the cognitive readiness skills of situational awareness, communication, and decision-making as well as captured instances where help was given by or sought from the facilitator (see Appendix 2.1). Correct damage control behaviors (e.g., checking a door for heat, applying the correct patching procedure, reporting that the water is rising above the deck plates) and key errors (e.g., opening a door or securing a valve without the required permission; allowing water to flood into a second, dry compartment; failing to secure a watertight door) were recorded. The time taken to perform key activities was also recorded in order to compare the speed at which tasks were completed.
- Bridge watch scenario—In the bridge watch task, recruits play different roles on the bridge. They must communicate with each other and, for some roles, with lookouts elsewhere on the ship. Those lookouts are reporting new ship or air contacts via the phone. Researchers were stationed near the phone-talkers and between the conning officer and facilitator and captured indicators for the cognitive readiness skill of communication (see Appendices 2.2 and 2.3). Researchers observing the phone-talkers listened to two of the stations (out of three possible) and recorded whether or not the phone-talker identified themselves and their target (e.g., port lookout), did a correct repeat back to the lookout, and concluded with "aye." Researchers observing the conning officer noted if they did correct repeat backs, as well as if the helm and lee helm did repeat backs, and reported task completion.
- Roving security watch scenario—The roving task requires recruits to navigate to the armory where they are assigned four different compartments they must find

and check to see if they are secure. The assigned compartments were distributed across multiple decks of the ship. The recruits are also instructed to watch out for and note any potential security or safety hazards they might encounter while on watch. Researchers observed the scenario briefing in the recruits' berthing compartment and then followed them throughout the course of the watch. Researchers captured indicators for the cognitive readiness skill of situational awareness, as well as captured instances where help was given by or sought from the facilitator (see Appendix 2.4). In particular, the time taken to navigate from the berthing compartment to the armory and subsequently to each compartment was recorded as well as the number of hazards encountered.

In developing the observation protocol for the team study, we followed a rigorous process to understand the BS21 environment, the performance of the recruits in that environment, and the skills and behaviors demonstrated in the BS21 scenarios. We performed an initial observation of BS21 in its usual operation. We analyzed and observed each target scenario to identify all decision points, typical recruit errors, and typical facilitator behavior. We performed a cognitive task analysis of the flooding control domain and identified the key skills and skill levels expected of the recruits at RTC. Based on a report on BS21 and discussions with BS21 instructors (HPC, 2008), we identified known issues with each scenario. From this information, we composed our observation forms. A different observation protocol was developed and tested for each scenario. As illustrated in Appendix 2, the observation protocols focused on objective measures of incidents of critical errors, communication events, whether key tasks were performed, time to perform tasks, number of incorrect attempts or inappropriate actions, number of facilitator interventions, and number of requests for assistance. These behaviors are consistent with the team-level readiness behaviors discussed by Bowers and Cannon-Bowers (this volume). In an "extra details" section of the survey, subjective observations were also captured. Researchers noted any behaviors they thought may be relevant to tasks but which were not specifically asked about in the survey. Due to accessibility limitations, we were only able to pilot our observation protocols on a no-game group. Previous research has demonstrated that raters are able to identify these behaviors reliably (Bowers et al., forthcoming).

17.5.2 Results

17.5.2.1 Situational Awareness

Groups were compared for differences in their awareness of their environment and current situation in that environment, including detection of novel cues as well as appropriate identification of expected cues. The ability to navigate effectively through the ship was captured as an element of situational awareness since it involved the ability to appropriately identify and interpret location cues as well as the ability to understand where they needed to be in order to handle the current situation or respond to changes in the situation.

In the roving security watch scenario, there was no significant difference in the number of hazards detected between the game experimental group and the control group. However, in the time required to find some of the assigned compartments, those in the game condition ($M = 1.29$ min, $SD = 1.30$) took significantly less time to navigate between berthing and the armory than those in the control condition ($M = 2.17$ min, $SD = 1.38$). Also, near the end of the task, those who played the game took significantly less time ($M = 1.36$, $SD = 1.13$) than the control group ($M = 2.29$, $SD = 1.54$) to navigate between the fourth compartment and the armory. These results do seem to suggest that initially, those in the game condition were more adept at being able to locate that first compartment.

In the flooding magazine compartment scenario, the group that played the game was significantly more likely to notice and report that the water rising above the deck plates was an important change in the situation (35 % of game groups reported vs. 13 % of no-game groups). However, the difference between groups was not significant statistically for the time between the water hitting the deck plates and the time it took them to report it.

17.5.2.2 Communication

The game and no-game groups were compared to assess differences in the frequency of communication errors. In the flooding magazine compartment scenario, there were no significant differences in the likelihood of requesting permission before entering the compartment. In the bridge watch scenario, participants that played the game were significantly less likely to make communication errors when talking to the bridge watch (14.59 % vs. 40.54 % error rate) and to the conning officer (4.77 % vs. 16.76 % error rate).

17.5.2.3 Decision-Making

Groups were compared for differences in the likelihood of making effective decisions or committing critical errors in the flooding magazine compartment scenario. More participants entered the flooding compartment without first checking the door for heat and pressure in the control group (31.8 %) compared to those in the game group (11.1 %). These differences approached statistical significance ($p = 0.051$).

The differences in the time between picking up the patch and then the wrench were significant. Those in the experimental condition were far faster ($M = 0.229$ min, $SD = 0.910$ min) than those in the control condition ($M = 1.000$ min, $SD = 1.523$ min). This might suggest that having played the game, the recruits were better aware of the fact that a wrench is needed for applying the patch and that perhaps they were more apt at the overall patching process. There were no significant differences in the time taken to repair the pipe.

With respect to whether the teams failed to secure the watertight doors behind them, those in the game condition made this error fewer times proportionally (13.3 % vs. 40.00 % for leaving the final door open); however, the cell sizes for these comparisons were not large enough for a statistical comparison.

We were interested in whether recruits committed the critical error of shutting off the water valve for the leaking pipe without first asking permission during the flooding scenario. Doing so would have cut off the water supply to recruits attempting to put out a fire on the deck above. In the VESSEL DCT game, recruits often made this mistake and immediately failed the level and had to view a cinematic of a teammate being burned alive as a consequence of their mistake. Out of the 15 in the control condition who addressed the water valve in some way, 33.3 % attempted to shut it off without first requesting permission, while out of the 23 in the game condition who addressed the valve, none (0 %) attempted to shut it off without first requesting permission. These differences were significant.

Moreover, those in the game condition took significantly less time to enter the flooding compartment and find the jubilee patch ($M = 2.17$ min, $SD = 1.83$) than those in the control condition ($M = 4.12$ min, $SD = 4.67$). Combined with the earlier result of a positive trend for checking the door for heat and 0 % likelihood of shutting the valve, these results indicate a general improvement in the speed with which the game group initiated the appropriate actions at the beginning of a time-sensitive task. In other words, the game group was faster at checking the door for heat, entering the compartment, identifying the leak problem, and locating and picking up the jubilee patch and wrench to begin repair.

17.5.2.4 Help Seeking

Groups were compared for differences in help-seeking behaviors and facilitator interventions. During the roving security watch, those in the game condition made significantly fewer requests for help ($M = 0.08$ requests, $SD = 0.28$) than those in the control condition ($M = 0.47$ requests, $SD = 0.83$). Also, the instructor intervened and provided assistance significantly more often to those in the control condition ($M = 2.08$, $SD = 3.23$) than to those in the game condition ($M = 1.33$, $SD = 1.61$). In the flooding magazine compartment scenario, the control group asked for help ($M = 1.57$ requests, $SD = 2.56$) significantly more often than those in the game group ($M = 0.30$ requests, $SD = 0.81$).

17.5.3 Discussion

A study was conducted to see if a brief exposure to a computer-game-based training experience could result in better performance in a complex damage control environment. In general, the results indicate that the game-based training was effective in improving many of these targeted skills. Groups that played the game were more significantly likely to communicate effectively and appropriately. However, this finding was contingent on type of scenario, i.e., not significant for the flooding magazine scenario but significant for the bridge watch scenario. It is presumed that the nature of the game's feedback allowed recruits to better understand the importance of each part of the required communication.

Results for situational awareness variables also provide support for the effectiveness of the game. The game group navigated more effectively. It is hypothesized that the game group was generally more aware of their objectives, methods, and surroundings, resulting in a lower need for help and an increased speed in initiating appropriate actions at the beginning of a task.

Groups that played the game were less significantly likely to make specific decision-making errors that could imperil the entire group. This is another area that was emphasized in the game, and the feedback was specifically designed to place the error in the larger context of the overall crew's work.

In particular, it seems that playing the game, complete with its graphic depiction of what happens when a recruit commits the critical error of securing a valve without permission, effectively trained recruits to not make this same mistake in the testing environment. This has a very interesting implication for the design of training games for areas such as damage control where mistakes can cost lives—perhaps forcing an individual to fail and be confronted with the catastrophic consequences in the game environment will effectively train them to not make the same mistake in the real world. Thus, the results are consistent with our theoretical position.

One should note that the test environment is a group-work environment and that there is an assessment challenge to objective observation of awareness across multiple individuals independently of how it is reflected in specific team actions, such as communication or decision-making.

There are several additional methodological limitations of the present research. Although a large number of recruits participated, they typically performed as groups, limiting the actual amount of usable data. Additionally, although senior personnel were instructed to act only as evaluators, they sometimes provided assistance to struggling groups to increase the value of the overall training experience, adding error to our measures.

These methodological limitations notwithstanding the data support the hypothesis that a brief contextual training experience, presented in a low-fidelity game to individuals, appears able to improve performance of teams in a much more complex, high-fidelity performance environment.

17.6 Comparison of Study Results

Table 17.1 summarizes the largest performance differences on measures of cognitive readiness skills of the game groups compared to the no-game groups in both the individual study and the team study. Each value in the table represents a significant performance difference on one measure for the corresponding cognitive readiness skill, and all values are oriented in the same direction. For example, if 40 % of the no-game group made a certain error and only 14 % of the game group made that error, then that is a 26 % higher performance for the game group; if the no-game group performed a task in 2.29 min and the game group performed it in 1.36, then that is a 41 % higher performance for the game group. In general, the results show that playing the VESSEL DCT produced significantly higher cognitive readiness for both individuals and teams but that the difference for teams was generally smaller. Additionally, while performance was significantly higher across almost many measures in the individual study, performance was significantly higher on only a few measures in the team study. It should be noted that performance measures in the individual study were based on testing in a controlled transfer scenario within the BS21 ship that was similar to a scenario in the game—a near-transfer task with few observational challenges—while measures in the team study were based on testing in an uncontrolled transfer scenario in the usual BS21 exercise—a far-transfer task with many observational challenges.

17.7 Team Assessment Challenges

The BS21 environment is a highly chaotic environment that provides a number of challenges to effective assessment of team performance. In conducting the team study, several anecdotal observations were made that call out some of the specific challenges of using a direct observation protocol for assessing the effects of individual cognitive readiness on team performance.

A single team member that is highly competent and cognitively ready can greatly skew indicators of team performance. This was anticipated based on pilot observations, and we tried to account for it through capturing requests for assistance and facilitator interventions. However, this was insufficient to impact our objective measures. Our anecdotal observation was that the effect of lone competent individuals exaggerated the observed competencies of the teams in the no-game group, while the game group appeared to have more competent individuals in their teams. Capturing these differences reliably in an observer-based protocol remains a challenge, however, and requires further investigation.

Anecdotally, there appeared to be a general increase in the amount of communication among team members in the game group, including both general chatter and task-specific discussion. For instance, in general, while a small number of members in a typical no-game team would vocalize their understanding of their current position relative to where they were supposed to be, many of the members in the game teams seemed to vocalize their understanding. Further, most of no-game group individuals who did vocalize seemed to use general terms (e.g., the numbers are going the wrong way, we need to turn around), while those in the game group seemed to use specific terms (e.g., we're heading aft, we need to turn around and head forward). There also appeared to be increased verbal collaboration in the game group, as compared to the no-game group. For instance, in the roving security watch scenario, the teams generally navigated the ship arranged in a line, with the leader at the front of the team. In many no-game teams, the members at the back of the team were almost totally silent and strained to hear what was being said at the front, while in several game teams, the members at the back of the team chattered softly about what they were doing and also relayed messages back and forth with teammates further up the line. These anecdotal observations suggest that a more complex observational protocol that captures incidental (i.e., not explicitly required for the task) verbal communications within the team may better detect cognitive readiness differences between teams.

17.8 Conclusions

We demonstrated that a game-based system for training individuals for enhanced cognitive readiness can produce higher team performance in situational awareness, communication, and decision-making. Training individuals to be more active, appropriate, accurate, and complete in their communication led to better team communication skills, and training individuals to be more aware of their surroundings and the consequences of their actions led to better team decision-making and situational awareness skills. While our results are positive, we believe that even stronger effects may be observed with enhanced observation protocols. Based on our experience, we make several recommendations for future efforts seeking to train to improve team performance and/or to assess team cognitive readiness in complex real-world contexts.

Training to improve communication skills appears to have a strong effect, particularly with novices, on several aspects of team performance and readiness. This suggests that communications training may be a high-value training intervention that should be considered early when preparing teams (or, if one is short on time to prepare a team, considered as the primary intervention).

For a team study, it is essential to do an initial analysis of how good teams differ from bad teams. Our observations were tied too closely to scenario-specific acts, not to the internal team dynamics. We had previously observed teams within the environment and piloted observational protocols with no-game subjects. However, the low level of internal team communications within the baseline groups did not lead us to consider additional measures that would more rigorously capture team internal communications. Further, it is very difficult to capture team data in real time using real-time, direct observational protocols. While we tried to capture some team communication behaviors, these were too challenging to capture effectively in a team of 8–11 in size with a single observer (while also moving around the environment in some scenarios). For instance, we often only could overhear the responses of one participant. In the absence of knowing the other part of the conversation, it was hard to do much more than record simple communication events.

Robustness measures are important when assessing team readiness. For instance, it is important to distinguish a team in which a small number are cognitively ready vs. a team in which a large number are, even if the overall team performance is similar. Capturing the source(s) of suggestions within the team and the strength and specificity of suggestions may be critical to objectively distinguishing performance among different groups.

While our results support the value of individual training for improving team performance, further study is needed in order to tease out exactly what the interplay is between individual performance and cognitive readiness and team performance and cognitive readiness.

Acknowledgments We would like to acknowledge the VESSEL team for their efforts in developing the game and supporting the studies, including Bruce Roberts, John Ostwald, Kerry Moffitt, Ray Tomlinson, and Todd Wright at Raytheon BBN Technologies; Curtiss Murphy, Chris Rodgers, Brad Anderegg, and David Guthrie at Alion Science and Technology; Katelyn Procci and Lucas Blair at the University of Central Florida; Kelly Pounds at IDEAS; and Susan Coleman and Ellen Menaker at Intelligent Decisions Systems, Inc. We also would like to thank Mr. John Drake and, previously, Dr. Rodney Chapman at the Naval Service Training Command and the staff at the RTC for their support and participation. The research presented in this chapter was conducted under Office of Naval Research contract number N00014-08-C-0030, under the management of Dr. Ray Perez. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the Office of Naval Research.

17.9 Appendix 1: Observer Form for Individual Study

The items in the observer form in this appendix are notated with the cognitive readiness skill they indicate as follows: D for decision-making, C for communication, and S for situational awareness.

Recruit Name_________________

Start Time: ___________________

AFTER REPORT TO DCC:

S1 Time to find 2-44-5-M: _________________
S2 Course Reversals:
S3 Wrong doors opened:
D1 Door left open:

Clear Timer

AT LOCKER

D2 ___ Equipped Helmet
D3 ___ Equipped boots
D4 ___ Equipped other gear

Start Timer

AT 2-44-5-M

S4 Time from repair locker:
C1 ___ Called DCC for permission
D5 ___ Entered without permission
D6 ___ Exited compartment
D7 ___ Closed door
C2 ___ Tried to repair without report

Clear Timer

Start Timer

BACK AT 2-44-5-M

17.10 Appendix 2: Observer Forms for Team Study

The items in the observer forms in this appendix are notated with the cognitive readiness skill they indicate as follows: D for decision-making, C for communication, and S for situational awareness. H denotes help given or sought.

17.10.1 Appendix 2.1: Flooding Magazine Compartment Scenario Observer Form

Facilitator & Division ______________________ # Recruits ________
Schedule time: ________

Facilitator Grade (Below-Avg, Avg, Above-Avg) ________
<u>**Flooding Control/Ammo Scenario**</u>
Time the Facilitator confirms they are ready to start Start Time: ________

D1	Heat Check or Pressure Check Door 1		________
D1a	Entered Door 1	Time:	________
C1	*Opened/Tried to Open Door 1 without permission*		________
D2	Door 1 Left Open		________
D3	*Tried to turn valve without permission*	Time:	________
D4	Got Jubilee Patch	Time:	________
D5	Got Wrench	Time:	________
C1	*Asked for permission to apply patch*	Time:	________
D6	Applied Patch Correctly (apply above & slide down)	Time:	________
D6a	Time started applying patch	Time:	________
D6b	Time started using wrench	Time:	________
D6c	Time stopped patching activities	Time:	________
D7	Heat Check or Pressure Check Door 2 (dividing door)		________
D7a	Entered Door 2	Time:	________
C2	*Opened/Tried to Open Door 2 without permission*		________
	Time water hit deck plates		
S1	Time they noted (to themselves) water rising	Time:	________
S2	Time rising water reported to facilitator	Time:	________
D8	Allowed water to start flooding into dry compartment	Time:	________
D9	Door 2 (dividing door) secured	Time:	________
D10	Heat Check or Pressure Check Door 3		________
C3	*Opened/Tried to Open Door 3 without permission*		________
D11	Left Door 3 Open (exit door)		________
H1	Number of Facilitator Interventions on Damage Control	Count:	________
H2	Number of Times Team asked for help on Damage Control	Count:	

Fac instruction 1 Cause: ______________________
Time: ________

Fac instruction 2 Cause: ______________________
Time: ________

Fac instruction 3 Cause: ______________________
Time: ________

Fac instruction 4 Cause: ______________________
Time: ________

17.10.2 Appendix 2.2: Bridge Watch Scenario: Bridge Talker Observer Form

All items in this form indicate the cognitive readiness skill of communication.

Extra details

How did they tackle the problem initially (stood around or seemed confused, multi-task, crowded around patch, etc)

____
____
____
____
____
____
____

Fac instruction 5	Cause:	__________
	Time:	__________
Fac instruction 6	Cause:	__________
	Time:	__________
Fac instruction 7	Cause:	__________
	Time:	__________
Fac instruction 8	Cause:	__________
	Time:	__________
Fac instruction 9	Cause:	__________
	Time:	__________
Fac instruction 10	Cause:	__________
	Time:	__________

Did they use the Ammo Loading Plan __________

H3 Number of Facilitator Interventions on Ammo Moving — Count: __________

H4 Number of Times Team asked for help on Ammo Moving — Count: __________

Time Facilitator ended mission and started debrief — Time: __________

17.10.3 Appendix 2.3: Bridge Watch Scenario: Conning Officer Observer Form

All items in this form indicate the cognitive readiness skill of communication.

Facilitator/Division ___________________ # Recruits ______________

Facilitator Grade (Below-Avg, Avg, Above-Avg) ______________

Bridge Watch - Phone talker

Instance	ID Target	ID Self	Repeat Back	Aye
1				
2				
3				
4				
5				
6				
7				
8				
9				
10				
11				
12				
13				
14				
15				

Notes:

17.10.4 Appendix 2.4: Roving Security Watch Scenario Observer Form

Facilitator & Division ________________________________ # Recruits ______________
Schedule Time ______________ ______________
Facilitator Grade (Below-Avg, Avg, Above-Avg) ______________

Bridge Watch - Conning Officer/Helm/Lee Helm

Instance	From C.O. to Fac	Fac To Helm (Helm repeat back)	Helm to Fac (Helm report task completion)	Fac to Lee Helm (Lee Repeat back)	Lee Helm to Fac (Lee report task completion)
1					
2					
3					
4					
5					
6					
7					
8					
9					
10					
11					
12					
13					
14					
15					
16					
17					
18					
19					
20					

Notes: ____

Bridge Watch - Conning Officer/Helm/Lee Helm

Instance	From C.O. to Fac	Fac To Helm (Helm repeat back)	Helm to Fac (Helm report task completion)	Fac to Lee Helm (Lee Repeat back)	Lee Helm to Fac (Lee report task completion)
21					
22					
23					
24					
25					
26					
27					

Facilitator & Division ________________________ # Recruits ____________________

Schedule Time:________________ ____________________

Facilitator Grade (Below-Avg, Avg, Above-Avg) ____________________

Roving Security Watch

S1a	Time Left Berthing Compartment	Start Time:	0:00
	Berthing Compartment ID	Tack-number	
	Armory Compartment ID	Tack-number	
S1b	Navigated to Armory	Time:	
S2a	Time left Armory	Time:	
	ID of compartment #1	Tack-number	
S2b	Time to Find compartment #1	Time:	
	ID of compartment #2	Tack-number	
S3	Time to Find compartment #2	Time:	
	ID of compartment #3	Tack-number	
S4	Time to Find compartment #3	Time:	
	ID of compartment #4	Tack-number	
S5	Time to Find compartment #4	Time:	
S6	*Went to wrong deck*	Count:	
S7	*Went to wrong compartment*	Count:	
S8a	Time told to return to Armory	Time:	
S8b	Time to return to Armory	Time:	
S9a	Time told to return to Berthing	Time:	
S9b	Time to return to Berthing	Time:	
S10	Number of hazards: (ask facilitator)	# Reported:	
		# encountered	
H1	Number of times Instructor Helped	Count:	
H2	Number of Times Team asked for help	Count:	

Notes:

References

Bowers, C., Hussain, T., Roberts, B., Cannon-Bowers, J., & Blair, L. (forthcoming). Preparing to practice: The use of a game-based simulation as a pre-training intervention for recruits, *Simulation and Gaming*.

Bransford, J. D., Brown, A. L., & Cocking, R. R. (Eds.). (1999). *How people learn: Brain, mind, experience, and school*. Washington, DC: National Academy Press.

Brydges, R., Carnahan, H., Rose, D., Rose, L., & Dubrowski, A. (2010). Coordinating progressive levels of simulation fidelity to maximize educational benefit. *Academic Medicine, 85*(5), 806–812.

Cannon-Bowers, J. A., & Bowers, C. (2009). Synthetic learning environments: On developing a science of simulation, games and virtual worlds for training. In S. W. Kozlowski & E. Salas (Eds.), *Learning, training, and development in organizations*. Boca Raton, FL: CRC Press.

Cannon-Bowers, J. A., & Bowers, C. (2011). Team development and functioning. In S. Zedeck (Ed.), *APA handbook of industrial and organizational psychology, Vol 1: Building and developing the organization* (pp. 597–650). Washington, DC: American Psychological Association.

Cartier, J. L., & Stewart, J. (2000). Teaching the nature of inquiry: Further development in a high school genetics curriculum. *Science Education, 9*, 247–267.

Chi, M. T. H., Glaser, R., & Farr, M. J. (Eds.). (1988). *The nature of expertise*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Christensen, C., Laerson, J. R., Jr., Abbott, A., Ardolino, A., Franz, T., & Pfeiffer, C. (2000). Decision making of clinical teams: Communication patterns and diagnostic error. *Medical Decision Making, 20*, 45–50.

De Giovanni, D., Roberts, T., & Norman, G. (2009). Relative effectiveness of high- versus low-fidelity simulation in learning heart sounds. *Medical Education, 43*(7), 661–668.

Diller, D., Roberts, B., Blankenship, S., & Nielsen, D. (2004). DARWARS Ambush!—Authoring lessons learned in a training game. *Proceedings of the 2004 Interservice/Industry Training, Simulation and Education Conference (I/ITSEC)*. Orlando, FL, December.

Fletcher, J. D. (2009). Education and training technology in the military. *Science, 323*, 72–75.

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science, 7*, 155–170.

Glaser, R. (1989). Expertise in learning: How do we think about instructional processes now that we have discovered knowledge structure? In D. Klahr & D. Kotosfky (Eds.), *Complex information processing: The impact of Herbert A. Simon* (pp. 269–282). Hillsdale, NJ: Lawrence Erlbaum Associates.

HPC (Human Performance Center), Naval Service Training Command Detachment. (2008). *Battle Stations 21 Performance Measurement Execution Analysis Report*. Prepared on February 15, 2008 for the Commander, Naval Service Training Command (NSTC).

Hussain, T., Feurzeig, W., Cannon-Bowers, J., Coleman, S., Koenig, A., Lee, J., et al. (2010). Development of game-based training systems: Lessons learned in an inter-disciplinary field in the making. In J. Cannon-Bowers & C. Bowers (Eds.), *Serious game design and development: Technologies for training and learning* (pp. 47–80). Hershey, PA: IGI Global.

Hussain, T. S., Roberts, B., Bowers, C., Cannon-Bowers, J. A., Menaker, E., Coleman, S., et al. (2012). Designing and developing effective training games for the US Navy (pp. 27–43). *M&S Journal*, *Spring*.

Hussain, T. S., Weil, S. A., Brunye, T., Sidman, J., Ferguson, W., & Alexander, A. L. (2008). Eliciting and evaluating teamwork within a multi-player game-based training environment. In H. F. O'Neil & R. S. Perez (Eds.), *Computer games and team and individual learning* (pp. 77–104). Amsterdam, Netherlands: Elsevier.

Jentsch, F., & Bowers, C. A. (1998). Evidence for the validity of PC-based simulations in studying aircrew coordination. *The International Journal of Aviation Psychology, 8*(3), 243–260. doi:10.1207/s15327108ijap0803_5.

Johnson, W. L. (2009). A simulation-based approach to training operational cultural competence. *In Proceedings of the 2009 MODSIM World Conference*. October 2009.

Johnson, W. L., Wang, N., & Wu, S. (2007). Experience with serious games for learning foreign languages and cultures. *Proceedings of the SimTecT Conference*. Australia.

Kennedy, D. M., & McComb, S. A. (2010). Merging internal and external processes: Examining the mental model convergence process through team communication. *Theoretical Issues in Ergonomics Science, 11*(4), 340–358.

Koenig, A. D., Lee, J., Iseli, M., & Wainess, R. (2009). A conceptual framework for assessing performance in games and simulations. *The Interservice/Industry Training, Simulation & Education Conference*. Orlando, FL.

Kolb, D. A. (1984). *Experiential learning: Experience as the source of learning and development*. Englewood Cliffs, NJ: Prentice-Hall.

Lassiter, D., Vaughn, J., Smaltz, V., & Morgan, B. (1990). Evaluation of the effects of training interventions on team performance. *Proceedings of the Human Factors and Ergonomics Society 24th Annual Meeting* (pp. 1372–2376). Santa Monica, CA: Human Factors and Ergonomics Society.

MacMillan, J., Entin, E., & Serfaty, D. (2004). Communication overhead: The hidden cost of team cognition. *Team cognition: Understanding the factors that drive process and performance* (pp. 61–82). Washington, DC: American Psychological Association.

Mayrath, M., Clarke-Midura, J. C., Robinson, D. H., & Schraw, G. (2011). *Technology-based assessments for 21st century skills: Theoretical and practical implications from modern research*. Charlotte, NC: Information Age Publishing.

Miller, M. (2012). *"Naval Inspector General Visits Navy's Only Boot Camp" naval service training command news*. Recruit Training Command Public Affairs, Jan 20. Retrieved March 19, 2012, from http://www1.netc.navy.mil/nstc/news_page_2012_01_20.asp.

Molesworth, B. R. C., Bennett, L., & Kehoe, E. J. (2011). Promoting learning, memory, and transfer in a time-constrained, high hazard environment. *Accident Analysis and Prevention, 43*(3), 932–938.

Morrison, J., & Fletcher, J. D. (2002). *Cognitive readiness*. IDA Paper P-3735. Alexandria, VA: Institute for Defense Analyses.

Murphy, C., Rodgers, C., Guthrie, D., & McDowell, P. (2008). Building training games with the Delta3D Simulation Core. *Tutorial given at the 2008 Interservice/Industry Training, Simulation and Education Conference*. Retrieved June 29, 2009, from http://www.delta3d.org/filemgmt_data/files/Tutorial_629_v1.1.pdf.

O'Neil, H. F., & Perez, R. S. (Eds.). (2008). *Computer games and team and individual learning*. Amsterdam, The Netherlands: Elsevier.

Patrashkova-Volzdoska, R., McComb, S., Green, S., & Compton, W. (2003). Examining a curvilinear relationship between communication frequency and team performance in cross-functional project teams. *IEEE Transactions on Engineering Management, 50*(3), 262–269.

Raghavan, K., Satoris, M. L., & Glaser, R. (1997). The impact of model-centered instruction on student learning: The area and volume units. *Journal of Computers in Mathematics and Science Teaching, 16*, 363–404.

Roberts, B., Diller, D., & Schmitt, D. (2006). Factors affecting the adoption of a training game. *Proceedings of the 2006 Interservice/Industry Training, Simulation and Education Conference*. Orlando, FL, December.

Rupp, A. A., Gushta, M., Mislevy, R. J., & Shaffer, D. W. (2010). Evidence-centered design of epistemic games: Measurement principles for complex learning environments. *Journal of Technology Learning and Assessment, 8*(4), 1–45.

Salas, E., Cannon-Bowers, J. A., & Johnston, J. (1997). How can you turn a team of experts into an expert team?: Emerging training strategies. In C. E. Zsambok & G. Klein (Eds.), *Naturalistic decision making* (pp. 359–370). Hillsdale, NJ: Lawrence Erlbaum Associates.

Salas, E., Fowlkes, J., Stout, R., Milanovich, D., & Prince, C. (1999). Does CRM training improve teamwork skills in the cockpit?: Two evaluation studies. *Human Factors, 41*(2), 326–343.

Scalese, R. J., Obeso, V. T., & Issenberg, B. S. (2007). Simulation technology for skills training and competency assessment in medical education. *Journal of General Internal Medicine, 23*(1), 46–49.
Smith, S. J., & Roehrs, C. J. (2009). High fidelity simulation: Factors correlated with nursing student satisfaction and self-confidence. *Nursing Education Perspectives, 30*(2), 74–78.
Stewart, J., Cartier, J. L., & Passmore, C. M. (2005). Developing an understanding through model-based inquiry. In M. S. Donovan & J. D. Bransford (Eds.), *How student learn: History, mathematics and science inquiry in the classroom* (pp. 515–565). Washington, DC: National Academies Press.
Swarz, J., Ousley, A., Magro, A., Rienzo, M., Burns, D., Lindsey, A. M., et al. (2010). CancerSpace: A simulation-based game for improving cancer-screening rates. *Computer Graphics and Applications, IEEE, 30*(1), 90–94.
Tobias, S., & Fletcher, J. D. (2011). *Computer games and instruction*. Charlotte, NC: Information Age Publishing, Inc.
Wears, R. L., Perry, S. J., Shapiro, M., Beach, C., Croskerry, P., & Behara, R. (2003). Shift changes among emergency physicians: Best of times, worst of times. *Proceedings of the 47th Annual Meeting of the Human Factors and Ergonomics Society* (pp. 1420–1423). Santa Monica, CA: HFES.
Yedidia, M. J., Gillespie, C. C., Kachur, E., Schwartz, M. D., Ockene, J., Chepaitis, A. E., et al. (2003). Effect of communications training on medical student performance. *Journal of the American Medical Association, 290*(9), 1157–1165.
Zimmerman, C., Raghavan, K., & Sartoris, M. L. (2003). The impact of MARS curriculum on students' ability to coordinate theory and evidence. *International Journal of Science Education, 25*, 1247–1271.

Index

H.F. O'Neil et al. (eds.), *Teaching and Measuring Cognitive Readiness*,
DOI 10.1007/978-1-4614-7579-8,

D

E