The Design, Development, and Application of Cognitive Task Analysis Methods: Studying and supporting macrocognition through a Naturalistic Decision Making lens.

by

Laura G. Militello

A synoptic report submitted in partial fulfilment for the requirements for the degree of
Doctor of Philosophy at the University of Central Lancashire

January 2026

RESEARCH STUDENT DECLARATION FORM

Type of Award **PhD by publication**

School **School of Health, Social Work, and Sport**

*Sections marked * delete as appropriate*

1. Concurrent registration for two or more academic awards

*I declare that while registered as a candidate for the research degree, I have not been a registered candidate or enrolled student for another award of the University or other academic or professional institution

2. Material submitted for another award

*I declare that no material contained in the synoptic report has been used in any other submission for an academic award and is solely my own work

3. Collaboration

Where a candidate's research programme is part of a collaborative project, the synoptic report must indicate in addition clearly the candidate's individual contribution and the extent of the collaboration. Please state below:

Not applicable

4. Use of a Proof-reader

*No proof-reading service was used in the compilation of this synoptic report.

Signature of Candidate Laura Militello

Print name: **Laura Militello**

Dedication

I would like to dedicate this synoptic report to Mark Sisson for all his encouragement and support.

Acknowledgements

I would like to take this opportunity to thank my colleagues. For anyone who has ever conducted research, it is clear that no idea, no manuscript, no research design truly belongs to an individual. I have been fortunate to be a part of so many collaborations with creative, brilliant researchers. It has been an honor to build on the work of those have come before us and set the stage for the methods and frameworks I helped form and apply over the course of my career.

In particular, I want to thank Pam Richards. I would not have considered working on a PhD this late in life without her encouragement. It has been wonderful to discuss methods, applications, and cultural differences in applied research between the U.K. and U.S. It has been a joy to interact with you in this way. I would also like to thank Sarah Hobbs who was willing to be a supervisor on this synoptic report even though cognitive task analysis is somewhat removed from her research area. Her patience, and genuine curiosity about naturalistic decision making has made the process so enjoyable. Many thanks to both Pam and Sarah for being so generous with their time and providing such valuable guidance throughout the process.

I would also like to thank my long-time colleagues, Gary Klein, Emilie Roth, Rob Hutton, Brian Moon, and Robert Hoffman. Yours are the voices in my head as I conduct research, deliver presentations, and write manuscripts.

Lastly, I would like to thank my current coworkers: Julie Diiulio, Katie Ernst, John Hendricks, Heather Kinsman, Rachel Morris, Oliver Smith, Christen Sushereba, Michelle Wang, and Steve Wolf. You all have been so kind and encouraging about this PhD project. And when it comes to conducting rigorous research in pursuit of applications that support people who make tough decisions in complex settings, I can imagine no better colleagues. You inspire me every day.

ABSTRACT

This synoptic report focuses on the design, development, and application of cognitive task analysis (CTA) within the context of naturalistic decision making. It reflects 30+ years of research, emphasizing my methodological contributions. After an introduction (Chapter 1) and discussion of the research context (Chapter 2), my contributions are described in terms of three themes. The first is the development of the applied cognitive task analysis (ACTA) suite of methods (Chapter 3). These methods were designed to make CTA methods accessible to practitioners who have not studied cognitive psychology or related fields. Since the 1998 publication of ACTA, the methods have been applied and adapted for use in a broad range of domains. The second theme is the use of CTA to inform training (Chapter 4). This chapter includes a case study describing the use of CTA to develop training for sepsis recognition, followed by a project to identify theory and evidence-based training design principles to inform the design of recognition skills training. The third theme is CTA to inform technology design (Chapter 5). This discussion highlights my efforts to disseminate and apply the decision-centered design framework to guide the design of decision support technologies, and the more recent contribution in the form of integrated cognitive analysis (ICA) to inform the design of human-autonomy teams.

The final chapter of this synoptic report synopsizes these themes and my contribution to knowledge and applied practice, as driven by societal challenges. For example, the ACTA methods were developed in response to a call for methods to understand decision making as more physical aspects of work were mechanized leaving humans to manage cognitive complexity. The resulting ACTA methods have been used to support significant contributions to research worldwide. Sepsis recognition training was needed to overcome limitations in physician training brought on by changes in rules for supervision. The resulting training has improved sepsis recognition and patient outcomes. As advanced autonomation and autonomy become reality, methods are needed to envision how the nature of work will change and how humans and machines will work together. The ICA has been used successfully to meet this need in military and space operations. Each of these contributions met a need at a point in time; however, the world continues to change requiring additional adaptations and new methods. The final chapter concludes by outlining my path forward as I continue to study challenging real-world problems, with a focus on meeting emerging challenges.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ACRONYMS

Acronym	Full Term
ACGME	Accreditation Council for Graduate Medical Education
ACTA	Applied Cognitive Task Analysis
AR	Augmented Reality
AWACS	Airborne Weapons and Control Systems
CCHMC	Cincinnati Children’s Hospital Medical Center
CDM	Critical Decision Method
CRC	Colorectal Cancer
CTA	Cognitive Task Analysis
DCD	Decision Centered Design
DM	Decision Making
DOA	Department of the Army
EHR	Electronic Health Record
ESM	Early Sepsis Management
HMT	Human Machine Teaming
ICA	Integrated Cognitive Analysis
NDM	Naturalistic Decision Making
SJDM	Society for Judgment and Decision Making
SSA	Screening and Surveillance App
SVT	Supraventricular Tachycardia
TALSAR	Training for Advanced Life Support in Austere Regions
TTT	Tourniquet Troubleshooting Training
U.K.	United Kingdom
U.S.	United States

The Design, Development, and Application of Cognitive Task Analysis Methods: Studying and supporting macrocognition through a Naturalistic Decision Making lens.

CHAPTER 1. INTRODUCTION TO THIS SYNOPTIC REPORT

1.1. Introduction

This synoptic report documents my contributions to the body of science, including the design and development of cognitive task analysis (CTA) methods to study decision making (DM) and expertise in complex environments. In addition to method creation, I also discuss contributions I have made relating to the use of CTA methods in applied research to design training and technologies to support skilled performance in military and healthcare domains.

1.2 Chapter structure

Chapter 2 provides background information about CTA and the Naturalistic Decision Making (NDM) movement that has been central to my research approach. Chapters 3, 4, and 5 each describe a challenge or gap, followed by a discussion of my contributions to the design and development of methods, the application of CTA methods to training design, and the application of CTA methods to technology design respectively. The final chapter provides a critical synopsis and addresses next steps in the refinement and application of CTA.

1.3 Research themes

Throughout my career, I have been influential in the development, dissemination, and application of CTA methods. This synoptic report summarizes contributions related to three research themes:

Theme 1: Design and development of CTA methods. Chapter 3 describes my leadership role in developing the Applied Cognitive Task Analysis (ACTA) suite of methods (Militello & Hutton, 1998), and efforts to describe (Hoffman & Militello, 2008), disseminate (Klein et al., 2017), and extend CTA methods (Militello, Salwei et al., 2023; Fitzgerald, et al., 2024). This chapter addresses the following research questions:

Theme 2: Applying findings from CTA to inform training. Chapter 4 describes use of CTA to inform high-impact training design for sepsis recognition in a Children's Hospital; the creation of the *Handbook of augmented reality training design principles* to guide software developers and instructional designers in creating evidence and theory based training using emerging technologies (Militello, Sushereba, Ramachandran, 2023); and innovative strategies for evaluating the impact of training (Sushereba & Militello et al., 2024; Sushereba, Fernandez et al., 2024). This chapter addresses the following research questions:

Theme 3: Applying findings from CTA to inform technology design. Chapter 5 describes my efforts to promulgate the decision-centered design (DCD) framework (Militello & Klein, 2013) as a strategy to use CTA to integrate invisible aspects of work into design; an application of DCD to design decision support for colorectal cancer screening (Militello et al., 2016; Militello et al., 2017); and recent efforts to integrate CTA with other cognitive engineering methods to better address the challenges of designing for an envisioned world with advanced automation and autonomy (Militello et al., 2019a; Militello et al., 2019b). This chapter addresses the following research questions:

Research questions related to theme 1 include: 1) *Can sophisticatedl methods such as the Critical Decision Method be streamlined to support practitioners in effectively conducting CTA without an extensive apprenticeship period of learning?* 2) *How can both theoretical perspective and procedures for conducting CTA be described and disseminated effectively?* 3) *Can we develop CTA workshops that effectively communicate both the perspective required to conduct CTA and the methods?* 4) *Can CTA methods be adapted for use in large-sample studies, and to inform the design of human-machine teams integrating advanced automation and autonomy?* Research questions related to theme two focus on applying CTA to training as instantiated in specific, applied projects: 5*) Can incident accounts elicited during CTA interviews be used to inform training scenarios with high cognitive fidelity for use in a sophisticated simulation center? 6) Does the resulting training increase sepsis recognition? 7) How to effectively measure the effects of training on context-specific cognitive performance?* Research questions related to theme 3 focus on contributions to technology design in the context of multi-dimensional teams: 8) *How to communicate to project managers, software developers, implementation scientists and others the contributions of CTA? 9) Could the DCD framework effectively guide the design of decision support in the context of colorectal cancer screening management? 10) Could CTA uncover barriers and facilitators, decision requirements, and search strategies across disparate health systems? 11) Can retrospective CTA methods be combined with other human factors methods to support designing for the envisioned world?*

Tables 1.1, 1.2, and 1.3 summarize the context, need and research questions related to each theme, as well as relevant publications, contributions to knowledge and practical impact. Appendix 1.1 includes a timeline of key publications related to these themes. A more complete list of publications can also be found at https://orcid.org/0000-0001-8445-5640.

It is important to note that all the contributions detailed in this synoptic report were the result of work accomplished by strong research teams. I highlight contributions for which I have had a meaningful role in the intellectual development of concepts; a leadership role in directing the research; and/or served as a core team member in implementing methods, interpreting data, and applying findings to a real-world problem.

1.4 Summary and Conclusions

This synoptic report is structured around three research themes and associated scientific and practical contributions. The next chapter sets the stage for discussing these research themes by providing background information about CTA methods and the NDM movement.

Table 1.1: Table 2.1: Summary of Theme 1 – Design and development of CTA methods. Note: Number of citations refers to SCOPUS tallies for all journal articles. Publications not tracked in SCOPUS use Google Scholar tallies instead and are indicated with an asterisk*.

Context, Need, and Research Questions	Relevant Publications, Citations, and Percent Contribution	Contribution to Knowledge and Practical Impact
Context: In the 1990s, most CTA methods required many hours or days of knowledge elicitation and many were poorly documented. ***Need***: Practical and rigorous methods for understanding cognitive challenges and skills in naturalistic settings. ***Research Questions***: Can sophisticated methods such as the Critical Decision Method be streamlined to support practitioners in effectively conducting CTA without an extensive apprenticeship period of learning?	**Militello, L.G**. & Hutton, R.J.B. (1998). Applied Cognitive Task Analysis (ACTA): A practitioner's toolkit for understanding cognitive task demands. *Ergonomics Special Issue: Task Analysis*, *41* (11), 1618 -1641. https://doi.org/10.1080/001401398186108 Citations: 461; Contribution Project management: 85% Conceptual: 50% Writing: 60%	***Novelty***: ACTA presented a streamlined suite of knowledge elicitation methods that could be conducted in less than 2 hours, representing an original contribution in terms of usability and pragmatic application of CTA methods. ***Original Findings***: ACTA methods leverage the strengths of more sophisticated techniques such as the Critical Decision Method, and provide scaffolding in the form of specific (tailorable) question probes and note-taking structures. An evaluation study demonstrated that after a 6-hour workshop introducing the ACTA techniques, graduate students were able to elicit relevant, domain-specific, accurate cognitive information that was translated into training materials. ***Practical Impact***: The original ACTA paper (Militello & Hutton, 1998) has been cited over 900 times. The methods have been used to study diverse domains by researchers and practitioners worldwide. Examples from my body of work include studies of software developer expertise (Alarcon et al

Table 1.1: Summary: of Theme 1 – Design and development of CTA methods continued. Note: Number of citations refers to SCOPUS tallies for all journal articles. Publications not tracked in SCOPUS use Google Scholar tallies instead and are indicated with an asterisk*.

Context, Need, and Research Questions	Relevant Publications, Citations, and Percent Contribution	Contribution to Knowledge and Practical Impact
		2017), internet navigation by blind users (Gorman et al., 2004), as well as military projects that are confidential. See Appendix 3.1 for a list of 27 published accounts of projects using ACTA in 11 different domains. ***Limitations***: ACTA is widely used; however, we have limited information about how it has been adapted and improved.
Context: In the U.S., many psychology, engineering, and management graduate programs require courses in quantitative methods such as inferential statistics, but few offer courses in qualitative methods. Similarly, most texts addressing the design of technology and training do not address the practical challenges of using qualitative methods to address applied research questions.	Hoffman, R. & **Militello, L.G**. (2008). *Perspectives on Cognitive Task Analysis: Historical Origins and Modern Communities of Practice*. Taylor and Francis. https://doi.org/10.4324/9780203809877 Citations: 37 Contribution Project management: 50%	***Novelty***: Workshops I designed with Gary Klein in the 1990s were some of the first attempts to make CTA widely accessible. Similarly, *Perspectives on CTA* is the only text I am aware of that describes the history of CTA and draws direct links to related methods and traditions. The demand for CTA workshops has continued for over three decades and they have expanded to include international audiences. We recently conducted a CTA workshop in New Zealand that was attended by people from the U.K., Australia, Korea, Singapore, and Estonia as part of the NDM meeting (Moon & Militello 2024). ***Original Findings***: We developed an interactive workshop that provided an overview of NDM and expertise, demonstration

Table 1.1: Summary of Theme 1 – Design and development of CTA methods continued. Note: Number of citations refers to SCOPUS tallies for all journal articles. Publications not tracked in SCOPUS use Google Scholar tallies instead and are indicated with an asterisk*.

Context, Need, and Research Questions	Relevant Publications, Citations, and Percent Contribution	Contribution to Knowledge and Practical Impact
Need: Strategies for disseminating CTA methods that are not typically taught in university courses. ***Research Questions***: How can both the theoretical perspective and procedure for conducting CTA be described and disseminated effectively? Can we develop CTA workshops that effectively communicate both the perspective required to conduct CTA and the methods?	Conceptual: 40% Writing: 45% Jonassen, D., **Militello, L.G.**, & Crandall, B. (1998). Critical Incident/Critical Decision Method. In D.H. Jonassen, M. Tessmer, and WH. Hannum (Eds.), *Task Analysis for Instructional Design*. Lawrence Erlbaum. https://doi.org/10.4324/9781410602657-28 Citations: 11* Contribution Project management: 10% Conceptual: 40% Writing: 80% Klein, G., **Militello, L.**, Dominguez, C., & Lintern, G.	interviews, practice opportunities, and coaching. The workshop was tailored to the time available (1/2 day to week-long). Content was tailored to the audience. Offerings include the Critical Decision Method, ACTA, qualitative data analysis, knowledge representation, decision-centered evaluation, and CTA applications. One version of the workshop is described in Klein et al. (2017). ***Practical Impact***: CTA methods have been published in textbooks that are used by graduate students and practitioners to learn the history of CTA and how the methods link to related methods (Hoffman & Militello, 2008; Jonassen et al., 1988). The CTA Institute I cofounded with Brian Moon and Rob Hutton offers self-study courses in how to conduct CTA. I have conducted live workshops at academic conferences, as well as for private companies and government agencies. I have been instrumental in making CTA methods accessible. My efforts have facilitated broad dissemination of CTA methods, which in turn has led to application, adaptation, and expansion of the methods for use in healthcare (Russ et al., 2017),

Table 1.1: Summary of Theme 1 – Design and development of CTA methods continued. Note: Number of citations refers to SCOPUS tallies for all journal articles. Publications not tracked in SCOPUS use Google Scholar tallies instead and are indicated with an.*

Context, Need, and Research Questions	Relevant Publications, Citations, and Percent Contribution	Contribution to Knowledge and Practical Impact
	(2017). A One-Day Workshop for Teaching Cognitive Systems Engineering Skills. In P.J. Smith & R.R. Hoffman (Eds.), *Cognitive Systems Engineering: The Future for a Changing World*. CRC Press. https://doi.org/10.1201/9781315572529-16 Citations: 0 Contribution Project management: 25% Conceptual: 25% Writing: 25%)	military (Hutchins et al., 2003), sport (Johnston & Morrison, 2016), power industry (Wang et al., 2024), aviation (Diiulio et al., 2017), and more. ***Limitations***: We have no mechanism to "certify" that people are qualified to conduct CTA. Some have criticized that by making the methods widely available in this way, anyone can claim they are conducting CTA. There are no mechanisms for rating the qualifications of a CTA practitioner or consensus measures for assessing the rigor of a specific CTA study.
Context: Early CTA methods were used to conduct small sample studies of individual expertise. More recently, there have been calls for methods to	**Militello, L. G.**, Salwei, M. E., Reale, C., Sushereba, C., Slagle, J. M., Gaba, D., Weinger, M.B., Rask, J., Faiman, J., Andreae, M.,	***Novelty***: A *scientific innovation* was the adaptation of one of the ACTA methods (Simulation Interview) for use in a large-scale study, conducted by interviewers with diverse backgrounds, in geographically dispersed study sites. This

Table 1.1: Summary of Theme 1 – Design and development of CTA methods continued. Note: Number of citations refers to SCOPUS tallies for all journal articles. Publications not tracked in SCOPUS use Google Scholar tallies instead and are indicated with an.*

Context, Need, and Research Questions	Relevant Publications, Citations, and Percent Contribution	Contribution to Knowledge and Practical Impact
understand systemic issues that increase complexity in the context of joint cognitive systems. The challenges humans must manage increasingly include managing advanced automation and autonomy, raising questions about how to best support human-machine teaming. ***Need***: Adaptions of CTA methods to meet emerging challenges **Research Questions**: Can CTA methods be adapted: • For use in large-sample studies? • To inform the design of human-machine teams integrating advanced automation and autonomy?	Burden, A.R., Ander, S.. (2023). Adapting Cognitive Task Analysis Methods for Use in a Large Sample Simulation Study of High-Risk Healthcare Events. *Journal of Cognitive Engineering and Decision Making*, *17*(4), 315-331. https://doi.org/10.1177/15553434231192283 Citations: 2 Contribution Project management: 60% Conceptual: 70% Writing: 85% Fitzgerald, G., Morris, R., **Militello, L.**, & Fletcher, J. (2024, September). Building Trust in Human-Machine Teaming for Autonomous Space Sensing.	journal article was given the best paper award by the *Journal of Cognitive Engineering and Decision Making* in 2023. A *practical innovation* was adaptation of one of the ACTA methods (Knowledge Audit) for use in designing to support human-machine teaming. The HMT Knowledge Audit (McDermott et al., 2018) was developed by colleagues at Mitre with a particular emphasis on designing for advanced automation and autonomy. Similarly, Borders et al (2017), developed the critical decision audit, adapting the knowledge audit to incorporate more aspects of the critical decision method. These innovations to the methods created by other researchers speak to the impact and reach ACTA has had in the 25 years since its publication. **Original Findings**: In Militello & Salwei et al. (2023), we found that it was possible to standardize the Simulation Interview even further for use in large-sample studies. In Fitzgerald et al., (2024) we applied the HMT Knowledge Audit to identify user interface design requirements for the design of advanced automation to support space domain awareness.

Table 1.1: Summary of Theme 1 – Design and development of CTA methods continued. Note: Number of citations refers to SCOPUS tallies for all journal articles. Publications not tracked in SCOPUS use Google Scholar tallies instead and are indicated with an.*

Context, Need, and Research Questions	Relevant Publications, Citations, and Percent Contribution	Contribution to Knowledge and Practical Impact
	Proceedings of the 25th Advanced Maui Optical and Space Surveillance (AMOS) Technologies Conference, 22-35. https://amostech.com/TechnicalPapers/2024/Machine-Learning-for-SDA/Fitzgerald.pdf https://amostech.com/TechnicalPapers/2024/Machine-Learning-for-SDA/Fitzgerald.pdf Citations: 0 Contribution Project management: 75% Conceptual: 60% Writing: 35%	**Practical Impact**: Identifying cognitive aspects of work to support design continues to be a challenge as the nature of work changes. These adaptations of CTA continue to fill that need as challenges shift from understanding the cognitive skills of an individual to designing for complex, distributed systems and integrating advanced automation and autonomy. **Limitations**: Use of CTA for large-scale studies requires tradeoffs. For example, increased standardization of the methods allows for broader application of the methods, but reduces the depth of the analysis. This standardized approach limits the opportunity for discovery of unexpected findings. Although applications of the HMT knowledge audit and critical decision audit suggest they are effective, no evaluations of the strengths and weaknesses of these adaptations have been conducted.

Table 1.2: Summary of Theme 2 - Applying findings from CTA to inform training. Note: Number of citations refers to SCOPUS tallies for all journal articles.

Context, Need, and Research Questions	Relevant Publications (# of citations)	Contribution to Knowledge and Practical Impact
Context: In 2003, the Accreditation Council for Graduate Medical Education (ACGME) in the U.S. implemented regulations that reduced the hours a medical resident may work, and increased the level of supervision residents receive. As a result, some residents complete their training without the acquiring the skills needed to recognize critical conditions such as sepsis. ***Need***: Strategies for developing training scenarios with cognitive fidelity. ***Research Questions***: Can incident accounts elicited during CTA	Patterson, M.D., **Militello, L.G.**, Bunger, A., Taylor, R.G., Wheeler, D.S., Klein, G., Geis, G.L. (2016). Leveraging the critical decision method to develop simulation-based training for early recognition of sepsis, *Journal of Cognitive Engineering and Decision Making, 10*(1), pp. 36-56. https://doi.org/10.1177/1555343416629520 **Citations**: 23 **Contribution** Project management: 25% Conceptual: 60%	***Novelty***: The sepsis training developed as part of this study was a first of its kind. ***Original Findings***: The perceptual skills identified in this study led training developers to incorporate cues into the simulation that had not been used previously. For example, hands and feet of manikins were chilled prior to the training session to simulate poor distal perfusion. The sensemaking skills identified in this study led training developers to reduce the support and resources available to resident physicians in the training so that they would have an opportunity to practice their sensemaking skills without supervision and scaffolding from faculty. ***Practical Impact***: A study showed that the residents exposed to the sepsis training scenarios performed better in early sepsis management (ESM) as measured by the ESM checklist score and the ESM global performance score (Geis et al., 2018).

Table 1.2: Summary of Theme 2 - Applying findings from CTA to inform training continued. Note: Number of citations refers to SCOPUS tallies for all journal articles.

Context, Need, and Research Questions	Relevant Publications (# of citations)	Contribution to Knowledge and Practical Impact
interviews be used to inform training scenarios with high cognitive fidelity for use in a sophisticated simulation center? Does the resulting training increase sepsis recognition?	Writing: 55%) Geis, G.L., Wheeler, D.S., Bunger, A., **Militello, L.G.**, Taylor, R.G., Bauer, J.P., Byczkowski, T.L., Kerrey, B.T., & Patterson, M.D. (2018). A validation argument for a simulation-based training course centered on assessment, recognition and early management of pediatric sepsis. *Simulation in Healthcare*, *13*(1), 16-26. *DOI:* 10.1097/SIH.0000000000000271 **Citations**: 11 **Contribution** Project management: 5% Conceptual: 15% Writing: 5%	Furthermore, in the 6 years since this study, faculty reported that they feel the residents are better at recognizing sepsis due to this training intervention and the training was expanded to include all medical personnel (i.e., resident physicians, faculty, nurses). ***Limitations***: The hospital offers two encouraging testimonials: 1) the hospital has continued to invest in this training, and 2) personnel report that more cases of sepsis are recognized early since the training was implemented, leading to better patient outcomes. However, the data to support these claims are not publicly available.

Table 1.2: Summary of Theme 2 - Applying findings from CTA to inform training continued. Note: Number of citations refers to SCOPUS tallies for all journal articles.

Context, Need, and Research Questions	Relevant Publications (# of citations)	Contribution to Knowledge and Practical Impact
Context: As technologies such as augmented reality (AR) have become affordable, the ability to present learners with photorealistic cues in an immersive environment offers exciting opportunities to develop innovative training. However, few resources are available to support training designers. Literature related to training recognition skills is published by different disciplines, and often in difficult to find outlets such as conference proceedings and edited books. ***Need***: Design principles for training recognition skills using AR. ***Research Questions***: How can NDM research and CTA methods inform training design principles for	**Militello, L.G.**, Sushereba, C. E., & Ramachandran, S. (2023). *Handbook of augmented reality training design principles*. Cambridge University Press. https://doi.org/10.1017/9781009216166.001 **Citations**: 11* **Contribution** Project management: 90% Conceptual: 80% Writing: 80%	***Novelty***: This is the first book bringing together AR training design principles based on theory and empirical evidence. ***Original Findings***: The *Handbook of Augmented Reality training Design* articulates 11 design principles that are particularly relevant to using AR technology to train recognition skills. Importantly, these training design principles are relevant regardless of the training technology used. They have relevance for anyone designing training for people who must quickly size up a situation and act. ***Practical Impact***: As the book is only recently published, it is hard to characterize the impact. Anecdotal reports and book reviews suggest the book fills a need and serves as a valuable resource for both researchers and training developers. ***Limitations***: Since the book was published, I have discovered literature related to error recognition and recovery training that suggest an additional principle, I wish I had included. Although, my intent was to be comprehensive in identifying relevant literatures, I suspect there are other relevant

Table 1.2: Summary of Theme 2 - Applying findings from CTA to inform training continued. Note: Number of citations refers to SCOPUS tallies for all journal articles.

Context, Need, and Research Questions	Relevant Publications (# of citations)	Contribution to Knowledge and Practical Impact
leveraging the strengths of emerging technologies such as AR?		theoretical and empirically-based principles that did not make it into the book.
Context**: Assessing the impact of training is a challenge, particularly for performance in the high-stakes domains NDM researchers study. Learner perceptions of training and test of knowledge are common, but strategies for effectively assessing transfer of training remain elusive. ***Need: Strategies for assessing transfer of training to novel situations ***Research Questions***: How to effectively measure the effects of training on context-specific cognitive performance?	**Militello, L.G.,** Sushereba, C.E., Cheng, T., Kaduce, M., Read, J.M., Wagner, E., Smith, O.W., Goolsby, C., Winner, J. (2023 August 14-17.) *Error recovery training for hemorrhage control* [Poster presentation]. Military Health Services Research Symposium. Kissimmee, FL. **Citations**: 0 **Contribution** Project management: 50% Conceptual: 60% Writing: 50%	***Novelty***: In addition to measuring learner perception and performance on a knowledge test, we developed 1) a set of snapshot scenarios to measure knowledge transfer to novel situations, 2) a physical performance measure to assess transfer from mental practice to physical performance (Sushereba et al., 2024a; Sushereba et al., 2024b)). ***Original Findings***: This combination of measures provided a more nuanced understanding of training efficacy and limitations than simple knowledge tests commonly used. ***Practical Impact***: These two pilot studies provided evidence compelling enough to motivate sponsors to fund larger scale evaluations of the proposed training application. ***Limitations***: The physical demonstration measures used in these pilot studies were not validated before implementation due to schedule and resource constraints. Both were unsuccessful in discriminating between control and

Table 1.2: Summary of Theme 2 - Applying findings from CTA to inform training continued. 3: Note: Number of citations refers to SCOPUS tallies for all journal articles.

Context, Need, and Research Questions	Relevant Publications (# of citations)	Contribution to Knowledge and Practical Impact
		experimental groups. With current follow-on funding, we will strengthen the physical demonstration measure and conduct a validation study to ensure the measure is sensitive enough to distinguish between experts and novices.

Table 1.3: Summary of Theme 3 -- Applying findings from CTA to inform technology design. Note: Number of citations refers to SCOPUS tallies for all journal articles.

Context, Need, and Research Questions	Relevant Publications (# of citations)	Contribution to Knowledge and Practical Impact
Context: As CTA practitioners began to get a voice at the table in design and development teams, it became clear that we needed to articulate a framework that would help project managers, software developers, implementation scientists, and others understand the contributions of CTA. To have an impact in design, it is critical that multi-disciplinary teams have a high-level of understanding of each contributor's methods and potential contributions. Because CTA is a relatively new approach, program managers may have difficulty	**Militello, L.G**. & Klein, G. (2013). Decision-Centered Design. In J.D. Lee & A. Kirlik (Eds.). *The Oxford Handbook of Cognitive Engineering* (pp. 261-271). Oxford: Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199757183.013.0016 **Citations**: 51 **Contribution** Project mngment: 50% Conceptual: 40% Writing: 75%	***Novelty***: Decision-Centered Design (DCD, Militello & Klein, 2013) is different from other cognitive engineering frameworks in that it emphasizes designing to support tough decisions – when technological support is most needed (and also has the opportunity to do the most harm if poorly designed). Others advocate for a more comprehensive approach, often beginning with routine operations. ***Original Findings***: DCD emphasizes the use of CTA methods to uncover expertise and decision requirements, elements of work that are commonly overlooked in more technology-centric engineering disciplines. DCD articulates how CTA findings play a key role in the design of technology that supports human-machine teaming. DCD explains how CTA findings inform evaluation strategies that assess how well a proposed design intervention supports cognitive activities

Table 1.3: Summary of Theme 3 -- Applying findings from CTA to inform technology design continued. Note: Number of citations refers to SCOPUS tallies for all journal articles.

Context, Need, and Research Questions	Relevant Publications (# of citations)	Contribution to Knowledge and Practical Impact
knowing how to integrate CTA into other design activities. ***Need***: A framework describing how CTA fits into the technology design process ***Research Questions***: How to communicate to project managers, software developers, implementation scientists and others the contributions of CTA?		***Practical Impact***: DCD has been applied to a broad range of design problems (Assadi e al, 2022; Gualtieri et al., 2012; Harle et al., 2019; Miller et al., 2003; Wang et al., 2024). ***Limitations***: The DCD framework implies a linear process. In reality, it is generally iterative. Furthermore, many projects exercise only a subset of the 5 steps. Descriptions of DCD may oversimplify this complex activity.
Context: This project took place during a period of rapid adoption of electronic health records (EHR) in the U.S. Government incentives for EHR adoption led to technology implementation that was out of sync with the work. Many clinicians found it difficult to find information stored	**Militello, L.G.**, Saleem. J.J., Borders, M.R., Sushereba, C.E., Haverkamp, D., Wolf, S.P., Doebbeling, B.N. (2016). Designing colorectal cancer screening decision support: A cognitive engineering enterprise. *Journal of Cognitive Engineering*	***Novelty***: This project resulted in a design strategy that was novel at the time– and is widely used today. Specifically, the prototype Screening and Surveillance Application (SSA) retrieved screening history and other information relevant to managing colorectal cancer screening from the EHR and displayed it in an at-a-glance visualization. Rather than redesigning each EHR interface (many were in use in the U.S.

Table 1.3: Summary of Theme 3 -- Applying findings from CTA to inform technology design continued. Note: Number of citations refers to SCOPUS tallies for all journal articles.

Context, Need, and Research Questions	Relevant Publications (# of citations)	Contribution to Knowledge and Practical Impact
in the EHR. Colorectal cancer screening is one aspect of preventive care that requires access to a patient's screening history and up-to-date knowledge of complex screening guidelines. As colorectal cancer is a particularly deadly but treatable cancer, the U.S. CDC was seeking strategies to support clinicians in managing screening for their patients. ***Need***: Decision support for primary care clinicians managing colorectal cancer screening for their patients **Research Questions**: Could the DCD framework effectively guide the design of decision support in this context?	*and Decision Making, 10*(1), 74-90. https://doi.org/10.1177/1555343416630875 **Citations:** 16 **Contribution** Project mngment: 85% Conceptual: 75% Writing: 75%	at the time of this study), the proposed SSA could be used as an adjunct to any EHR (Militello et al., 2016). ***Original Findings***: An evaluation of the SSA suggested that the SSA was effective in supporting clinicians in managing CRC screening for their patients. Participants were able to answer questions about CRC related patient data accurately using the SSA. The SSA required less workload than the EHR alone to find CRC related patient information. Participants rated the SSA as highly usable and useful. ***Practical Impact***: This project demonstrated the power of CTA and DCD for identifying cognitive requirements to inform and guide design. User experience designers and software developers reported that the decision requirements table provided more useful design guidance than many other types of requirements documents. They found it helpful to understand the cognitive challenges the SSA was intended to support.

Table 1.3: Summary of Theme 3 -- Applying findings from CTA to inform technology design continued. Note: Number of citations refers to SCOPUS tallies for all journal articles.

Context, Need, and Research Questions	Relevant Publications (# of citations)	Contribution to Knowledge and Practical Impact
Could CTA uncover barriers and facilitators, decision requirements, and search strategies across disparate health systems?		***Limitations***: We were unable to transition the SSA from a prototype to a commercial product. At the time of the study, the Veterans Health Administration Medical Center that served as a study site requested the SSA be implemented in their health system. We were not able to overcome the bureaucratic barriers to implementation.
Context: Advances in automation and autonomy inspire optimism for reducing workload, increasing safety, and improving productivity. However, they also add new complexities and create new design challenges. ***Need***: Methods to support designing for the envisioned world. ***Research questions***: Can retrospective CTA methods be combined with other	Roth, E.M., Sushereba, C., **Militello, L.G.**, Diiulio, J., Ernst, K. (2019). Function Allocation Considerations in the Era of Human Autonomy Teaming. *Journal of Cognitive Engineering and Decision Making* 13(4), pp. 199–220. https://doi.org/10.1177/1555343419878038 **Citations**: 66	***Novelty***: As the U.S. Army funded the rapid development of rotorcraft that would fly further and faster than existing helicopters, I worked with a team to design crewing configurations. This envisioned world problem included not-yet-available advanced automation and autonomy and a near peer, nation state adversary. ***Original Findings***: We proposed and exercised the Integrated Cognitive Analysis (ICA; Militello, et al., 2019a; 2019b) for human machine teaming methodology as part of this project. Using this methodology, we made recommendations regarding what types of missions could safely and effectively be accomplished with reduced crewing given anticipated

Table 1.3: Summary of Theme 3 -- Applying findings from CTA to inform technology design continued. Note: Number of citations refers to SCOPUS tallies for all journal articles.

Context, Need, and Research Questions	**Relevant Publications (# of citations)**	**Contribution to Knowledge and Practical Impact**
human factors methods to support designing for the envisioned world?	**Contribution** Project mngment: 35% Conceptual: 35% Writing: 20%	technology capabilities. (Note: The specific findings from this work are confidential.) ***Practical Impact***: The ICA methodology was later used by NASA to make crewing recommendations for Mars missions. I am currently leading a project in which we are using the ICA methodology to design anomaly response systems on future smart habitats in space. ***Limitations***: The ICA methodology is resource and labor intensive, so it may not be practical in all cases. ICA requires a diverse set of skills including cognitive task analysis, design, discrete event modeling, and cognitive work analysis. Further, it requires access to stakeholders, domain experts, and technology developers.

CHAPTER 2. COGNITIVE TASK ANALYSIS AND NATURALISTIC DECISION MAKING

2.1 Introduction

Although there has long been an interest in studying decision-making dating as far back as the early psychotechnics in the 1880s (Hoffman & Militello, 2008), it is only in recent decades that CTA methods were developed to support applied research. This movement represented a departure from established methods that had limited practical value for emerging challenges such as designing decision support technologies and training personnel to make decisions in complex, real world, high stakes environments. For example, the Society for Judgment and Decision Making (SJDM) was founded in 1980, bringing together scientists who use experimental methods to study decision-making. These studies tend to use carefully defined tasks with university students (novices) to explore the impact of specific variables. An important contribution of this approach has been to provide insight into how to design situations in which humans can be easily fooled (Dale, 2015; Kahneman, 2011), leading to prescriptive models of DM intended to reduce bias and increase the rationality of DM. The resulting prescriptive models have been codified into protocols such as the military DM process (Department of the Army (DOA), 2019) and multi attribute utility theory (Dyer, 2005) in which decision makers are encouraged to identify criteria for success, generate multiple options, and then compare the options based on success criteria. These approaches are most valuable in situations that include a well-defined problem, time to conduct analyses and gather information, and a need to justify the decision to leadership. However, there is limited evidence that people use these formal methods of DM even when the conditions are right (Klein, 2009). As they were applied more broadly, it became clear that these analytic approaches quickly fall apart under time pressure when there is no time to gather information to conduct analyses, and can actually degrade performance. Naturalistic Decision Making (NDM) methods such as CTA were designed to overcome these limitations.

The following sections include brief context for my research journey, a discussion of CTA methods in the context of the NDM movement, and overview of the evolution of CTA.

2.2 Contextualizing the researcher journey

I joined Klein Associates in 1991, a small company led by the now world-renowned scientist Gary Klein. This experience early in my career laid the solid foundation for my research journey. I was involved in a number of ongoing research projects that provided an opportunity to apply and refine the Critical Decision Method (CDM). Such projects included interviewing United States (U.S.) Air Force F-15 pilots about dog fighting tactics, female heads of household about laundry stain removal strategies, Airborne Weapons and Control Systems (AWACS) weapons directors about directing aircraft in crowded airspace during military conflict, skilled musicians about one-on-one training, and critical care nurses about recognizing necrotizing enterocolitis in premature infants. At this time, in addition to applying Cognitive Task Analysis (CTA) to real-world problems, we were exploring how to teach CTA methods. Although we adapted them for each project, it seemed important to articulate core components of the methods so that they would be repeatable. We were also discussing what to call them. Was what we were doing CTA? If so, did CDM alone constitute CTA, or was there more to CTA? Being a part of these discussions provided valuable lessons for the later design of ACTA (Militello & Hutton, 1998), a key focus of this synoptic report.

2.3 Naturalistic Decision Making

The NDM movement began in 1989 with a meeting of like-minded researchers who coined the term NDM. The term "naturalistic" was intended to represent a focus on developing methods for studying DM in situ with an emphasis on skilled performance. NDM researchers focused primarily on domains characterized by high stakes and time pressure. They recognized that decisions in these contexts often involve multiple players and are generally influenced by organizational constraints. Goals are often vague and decision makers must use their experience to manage uncertainty (Klein et al., 1993). This characterization of DM in high stakes, dynamic settings was a novel conceptualization at the time. Although this nascent community was not describing their methods as CTA, they began to describe the interview and observation methods needed to study DM in these complex settings. (Klein et al., 1993).

The CTA methods that arose from NDM research were in direct response to the limitations uncovered when relying solely on an experimental approach to studying and supporting DM. Despite a long history of decision-making research, critical questions remained: How do skilled performers make decisions, particularly when they do not have the luxury of time to gather information and conduct a careful analysis? How do they manage when they have to make a decision in the face of uncertainty? How does their experience help (or hinder) them? CTA methods were designed to answer these questions. The goal of CTA was to study cognition in real-world contexts, requiring a departure from the controlled experiments traditionally used to study DM.

The term CTA highlights three important components (Klein & Militello, 2001). *Cognitive* refers a focus on how people think. The NDM community often uses the term *macrocognition* to refer to complex cognitive skills such as DM, problem solving, coordinating, and sensemaking as the focus of study. This is distinct from the study of *microcognition*, or laboratory-based studies of components of cognition such as memory or attention. *Task* refers to the fact that the methods are designed to study actual work. The focus of study is on situations in which the decision maker generally cares deeply about the outcome and brings relevant experience to the situation. *Analysis* refers to the use of systematic review of interview and observation data to obtain insights. In short, CTA can be defined as “a set of methods for identifying cognitive skills, or mental demands, needed to perform a task proficiently” (Militello & Hutton, p. 1618).

2.4 The evolution of CTA

The phrase CTA first emerged in the late 1970s in discussions of how to analyze cognitive aspects of work and how to help novices think more like experts (Glaser & Resnick, 1972; Resnick, 1976). At the time, a societal shift in the nature of work was occurring. Physical, repetitive tasks were increasingly taken over by technology, leaving more cognitively complex tasks to humans. In this context, the limitations of traditional task analysis methods that focused primarily on observable behaviors became highly salient. Thus, there was a call for methods to better understand the cognitive (and often invisible) aspects of work. Related methods evolved somewhat independently across different communities of practice (See Yates & Feldon, 2011 for a detailed review and classification of over 100 CTA methods described in the scientific literature). Traditions that have influenced NDM methods are summarized in Hoffman & Militello (2008). Over the last three decades, the NDM community has been instrumental in integrating models and methods from different research communities to develop and refine CTA methods.

In 1998, Klein and colleagues published a paper describing the CDM for eliciting knowledge. They described the method as including “probes that elicit aspects of expertise such as the basis for making perceptual discriminations, conceptual discriminations, typicality judgments, and critical cues.“ (Klein et al., 1998, p. 642). This became the basis for the most well-established CTA method today.

As use of CDM became increasingly widespread, Hoffman et al (1998) explored issues of reliability, adaptability, and validity of CTA with a focus on the CDM nearly 30 years ago.

Regarding reliability of the knowledge elicitation component of CTA, they reviewed studies in which interviewees reported incidents in the same way 3 and 5 months after an initial interview (Taynor et al., 1987). They also highlighted studies in which reliability of analysis of CTA data was established using measures of inter-rater agreement (Taynor et al., 1987; Kaempf et al., 1992). With regard to adaptability, they offer several examples in which CTA methods were tailored to specific research questions and constraints (Klein & Thordsen, 1991; Thordsen et al, 1992). With regard to validity, they focused on evidence of content validity, citing a study in which findings from CDM interviews were judged to be accurate, useful, and specific by an independent sample of nurses. Furthermore, the findings revealed aspects of nursing expertise not captured in existing training materials. (Crandall & Calderwood, 1989; Crandall & Gamblion, 1991; Crandall & Getchel-Reiter, 1993). Although traditional measures of reliability and validity derived from experimental studies are difficult to apply to qualitative methods, the NDM community continues to explore how to characterize rigor in a meaningful way (Klein et al., 2023).

2.5 Summary and Conclusions

CTA has emerged and matured over the last 40 years as the nature of work changed in the context of the evolving information age. The increasing cognitive complexity of the workplace called for new approaches to training and tools to support workers in high stakes domains characterized by time pressure, uncertainty, and vague goals, multiple players, and organizational constraints. CTA methods evolved in response to this need, offering new strategies for studying DM and other macrocognitive skills in context. Related methods have arisen across many research traditions. I have been part of the NDM movement that has actively integrated, refined, adapted, and evaluated the strengths and limits of CTA methods.

CHAPTER 3. DESIGN AND DEVELOPMENT OF APPLIED COGNITIVE TASK ANALYSIS

3.1 Introduction

This chapter focuses on my role in the design and development of ACTA. The U.S. Navy articulated a need for streamlined CTA methods that could be used by instructional system designers and instructors to more systematically incorporate cognitive challenges and effective strategies into the curriculum. ACTA (Militello & Hutton,1998) was developed to meet this need.

This chapter includes an overview of the ACTA methods (Section 3.3), including how they were developed (Section 3.2) and evaluated (Section 3.4), and an epilogue describing the influence of the methods over the las 25+ years (Section 3.5). The CDM was the prevailing CTA method in use at Klein Associates at the time. It provided a powerful tool for helping experts describe these difficult to articulate aspects of their performance, but it required practice and skill to conduct CDM interviews. I led a team of researchers through a 2.5 year project to develop streamlined CTA methods that would leverage some of the strengths of the CDM and also be accessible to practitioners. The ACTA evaluation was one of very few rigorous studies of the effectiveness of CTA methods at that time. Section 3.7 contains a broader discussion of my contributions disseminating ACTA and other CTA methods, as well as adapting and expanding CTA to support a range of research applications (Section 3.6). The chapter concludes with a brief summary and conclusions for the chapter (Section 3.7).

3.2 ACTA Methods Development

I began the project by reflecting on the more challenging aspects of conducting CDM interviews. These challenges included: quickly getting up to speed in a new job domain, eliciting an incident, exploring an incident in depth, managing the interview process, asking cognitive probes at appropriate times, taking notes, and managing qualitative data. We hosted a series of team meetings to brainstorm different strategies for addressing these challenges. We tried different interview strategies on each other, family, and friends until we had a draft set of methods. I documented each interview strategy and then tried them out on Navy personnel and firefighters, continuing to refine the probes and discuss their practicality within the team.

To ensure that the methods were accessible to practitioners, we developed a workshop and trained Navy personnel to use the methods. I obtained feedback from workshop participants and further refined the methods.

3.3 Results: The ACTA suite of methods

The ACTA methods include the Task Diagram, the Knowledge Audit, and the Simulation Interview, each with a particular focus (Figure 3.1). This suite of complementary methods can be used together or independently. They are not steps in a sequential process. Researchers are encouraged to emphasize different ACTA methods depending on the needs of each project.

Task Diagram → Overview
Knowledge Audit → Expertise
Simulation Interview → Incident based

Figure 3.1 : Overview of ACTA Suite of Methods

3.3.1 The Task Diagram for getting up to speed in a new domain

Most CTA practitioners work in many different domains; therefore, obtaining an overview of the domain prior to conducting interviews is an important step in tailoring the interview guide to the current domain. This preparation phase may take several weeks as researchers conduct literature reviews, speak with stakeholders, review training manuals, etc. to learn about the major tasks involved in the domain.

The Task Diagram was developed to streamline this preparation phase. If the project schedule does not allow for lengthy preparation, the knowledge audit can be used as a strategy to quickly orient to a new domain. The interviewer asks the interviewee to break down the task of interest into 3-6 steps, and then identify which of the steps described require difficult cognitive skills. The interviewer draws the steps in sequential order on a whiteboard and circles those that are complex (Figure 3.2). This interview generally takes 20 minutes or less and the resulting overview can serve as a roadmap for future interviews. The task diagram might be used as an initial step in bootstrapping an understanding of a new domain (Potter et al., 2000).

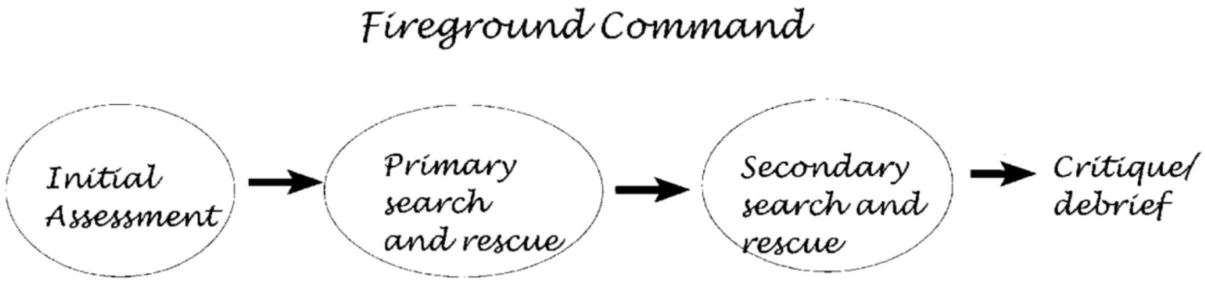
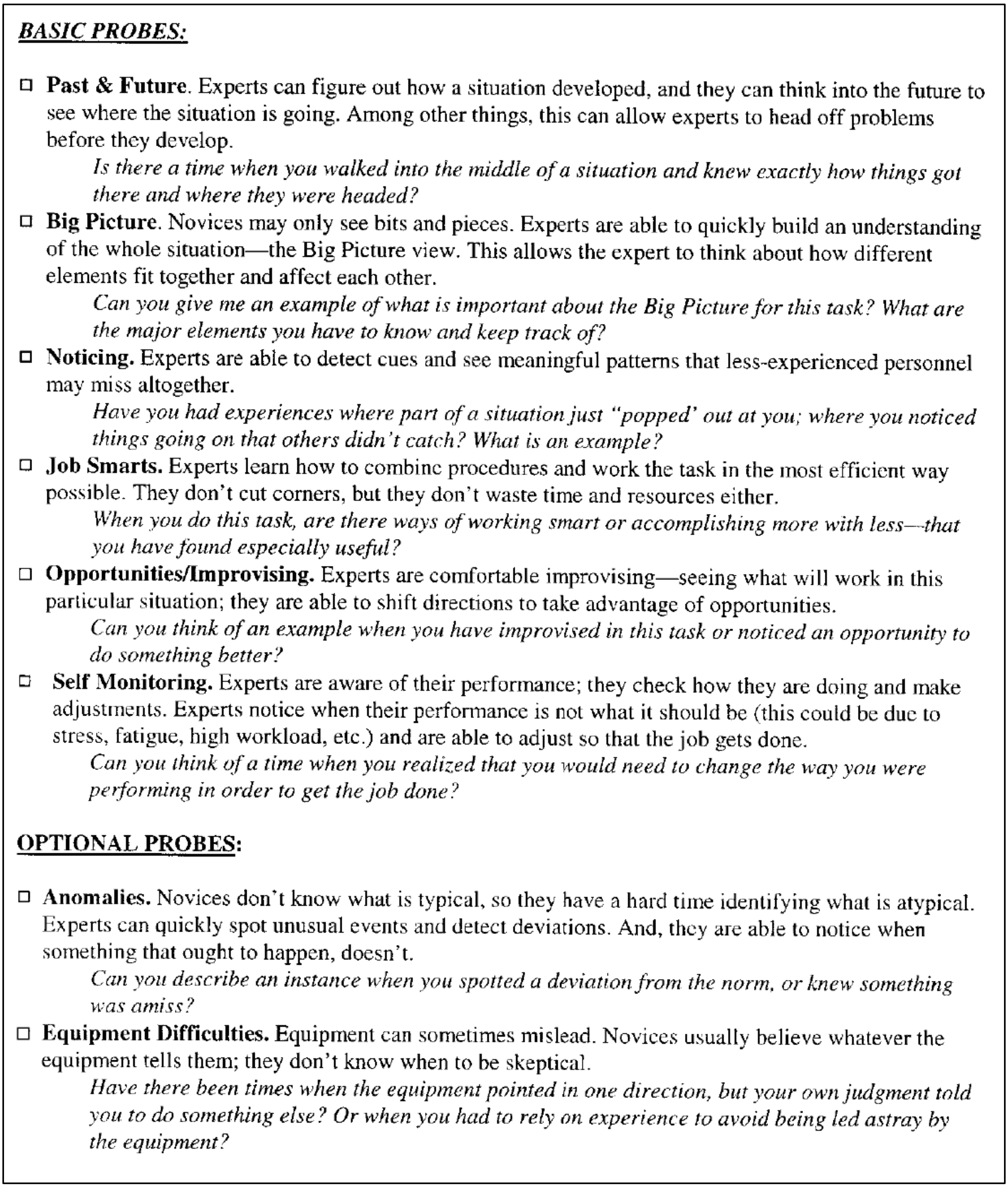


Figure 3.2: Sample Task Diagram (Militello & Hutton, 1998, p. 1621)

3.3.1 The Knowledge Audit for eliciting examples

To explain how we incorporated elements of the CDM (Klein et al, 1989) into ACTA, I first describe CDM. The CDM is a retrospective interview technique in which interviewees are asked to recall a challenging incident. The interviewer spends up to two hours exploring the incident to understand how it unfolded from the first-person perspective of the interviewee, including what cues were noticed, how cues were interpreted to make sense of the situation, and what actions were taken. The technique of grounding the interview in a lived experience is an important strategy for eliciting accurate, detailed accounts that are often difficult for experts to articulate without this type of support. However, the first step of the CDM eliciting a challenging incident can be particularly challenging for a new interviewer.

The Knowledge Audit was designed to elicit aspects of expertise without the difficult task of working with the interviewee to identify a rich, challenging incident that would provide insights relevant to study goals. We developed a series of probes (Figure 3.3) to elicit short examples of aspects of expertise rather than an in-depth account of a single incident. Drawing from the expertise literature, we created probes addressing specific aspects of expertise including diagnosing and predicting, situation awareness, perceptual skills, developing and knowing when to apply tricks of the trade, improvising, metacognition, recognizing anomalies, and compensating for equipment limitations. We refined the language for each probe, removing jargon from the scientific literature so that the probes would resonate with interviewees across many domains. The interviewer is encouraged to use probes to elicit examples and then follow up with questions about cues and strategies used, and what makes this aspect of work difficult.



BASIC PROBES:

☐ **Past & Future**. Experts can figure out how a situation developed, and they can think into the future to see where the situation is going. Among other things, this can allow experts to head off problems before they develop.
Is there a time when you walked into the middle of a situation and knew exactly how things got there and where they were headed?

☐ **Big Picture**. Novices may only see bits and pieces. Experts are able to quickly build an understanding of the whole situation—the Big Picture view. This allows the expert to think about how different elements fit together and affect each other.
Can you give me an example of what is important about the Big Picture for this task? What are the major elements you have to know and keep track of?

☐ **Noticing.** Experts are able to detect cues and see meaningful patterns that less-experienced personnel may miss altogether.
Have you had experiences where part of a situation just "popped' out at you; where you noticed things going on that others didn't catch? What is an example?

☐ **Job Smarts.** Experts learn how to combine procedures and work the task in the most efficient way possible. They don't cut corners, but they don't waste time and resources either.
When you do this task, are there ways of working smart or accomplishing more with less—that you have found especially useful?

☐ **Opportunities/Improvising.** Experts are comfortable improvising—seeing what will work in this particular situation; they are able to shift directions to take advantage of opportunities.
Can you think of an example when you have improvised in this task or noticed an opportunity to do something better?

☐ **Self Monitoring.** Experts are aware of their performance; they check how they are doing and make adjustments. Experts notice when their performance is not what it should be (this could be due to stress, fatigue, high workload, etc.) and are able to adjust so that the job gets done.
Can you think of a time when you realized that you would need to change the way you were performing in order to get the job done?

OPTIONAL PROBES:

☐ **Anomalies.** Novices don't know what is typical, so they have a hard time identifying what is atypical. Experts can quickly spot unusual events and detect deviations. And, they are able to notice when something that ought to happen, doesn't.
Can you describe an instance when you spotted a deviation from the norm, or knew something was amiss?

☐ **Equipment Difficulties.** Equipment can sometimes mislead. Novices usually believe whatever the equipment tells them; they don't know when to be skeptical.
Have there been times when the equipment pointed in one direction, but your own judgment told you to do something else? Or when you had to rely on experience to avoid being led astray by the equipment?

Figure 3.3: Knowledge Audit Probes (Militello & Hutton, 1998 p. 1622)

Notes are recorded on a shared display such as a whiteboard, using the matrix depicted in Figure 3.4. The short vignette examples that are elicited can be candidate topics for enhancing curricula to address cognitively complex aspects of the job.

Aspects of expertise	Cues and strategies	Why difficult?
Past and future e.g. Explosion in office strip; search the office areas rather than source of explosion	Material safety data sheets (MSDS) tell you that explosion in area of dangerous chemicals and information about chemicals Start where most likely to find victims and own safety considerations	Novice would be trained to start at source and work out May not look at MSDS, to find potential source of explosion, and account for where people are most likely to be
Big picture Big picture includes source of hazard, potential location of victims, ingress/egress routes, other hazards	Senses, communication with others, building owners, MSDS, building pre-plans	Novice gets tunnel vision, focuses on one thing, e.g. victims
Noticing Breathing sounds of victims	Both you and partner stop, hold your breath, and listen Listen for crying, talking to themselves, victims knocking things over	Noise from own breathing in apparatus, fire noises Don't know what kinds of sounds to listen for

Figure 3.4: Sample Knowledge Audit notetaking form (Militello & Hutton, 1998, p.1623)

3.3.4 The Simulation Interview for in-depth exploration of an incident

For the Simulation Interview, we focused on the part of the CDM in which the interviewer elicits an incident and creates a timeline before administering a series of cognitive probes. Often interviewers struggle to elicit a timeline before moving on to more in-depth probes, sometimes resulting in a jumbled account of the incident.

To streamline the in-depth exploration, we created the Simulation Interview. This interview technique begins with a simulated incident that the interviewee works through. The simulation can be low-tech (e.g., text-based), or use a high-tech simulator. The key is to have the interviewee experience a challenging incident that is relevant to the topic of study. Supplying a scenario in this way eliminates the need to elicit an incident and build a timeline.

The interview begins with the interviewee working through the scenario. The interviewer then asks the interviewee to identify major events, including judgments and decisions they made as they managed the scenario. Each event is probed for situation assessment, critical cues, and potential errors associated with that portion of the scenario, using CDM-type probes. The resulting account provides insight into what cues the interviewee noticed, how they made sense of the situation, actions they took or considered and rejected, and where they would expect a less experienced person struggle with the incident. Notes are captured in a shared display such as a whiteboard using the matrix depicted in Figure 3.5.

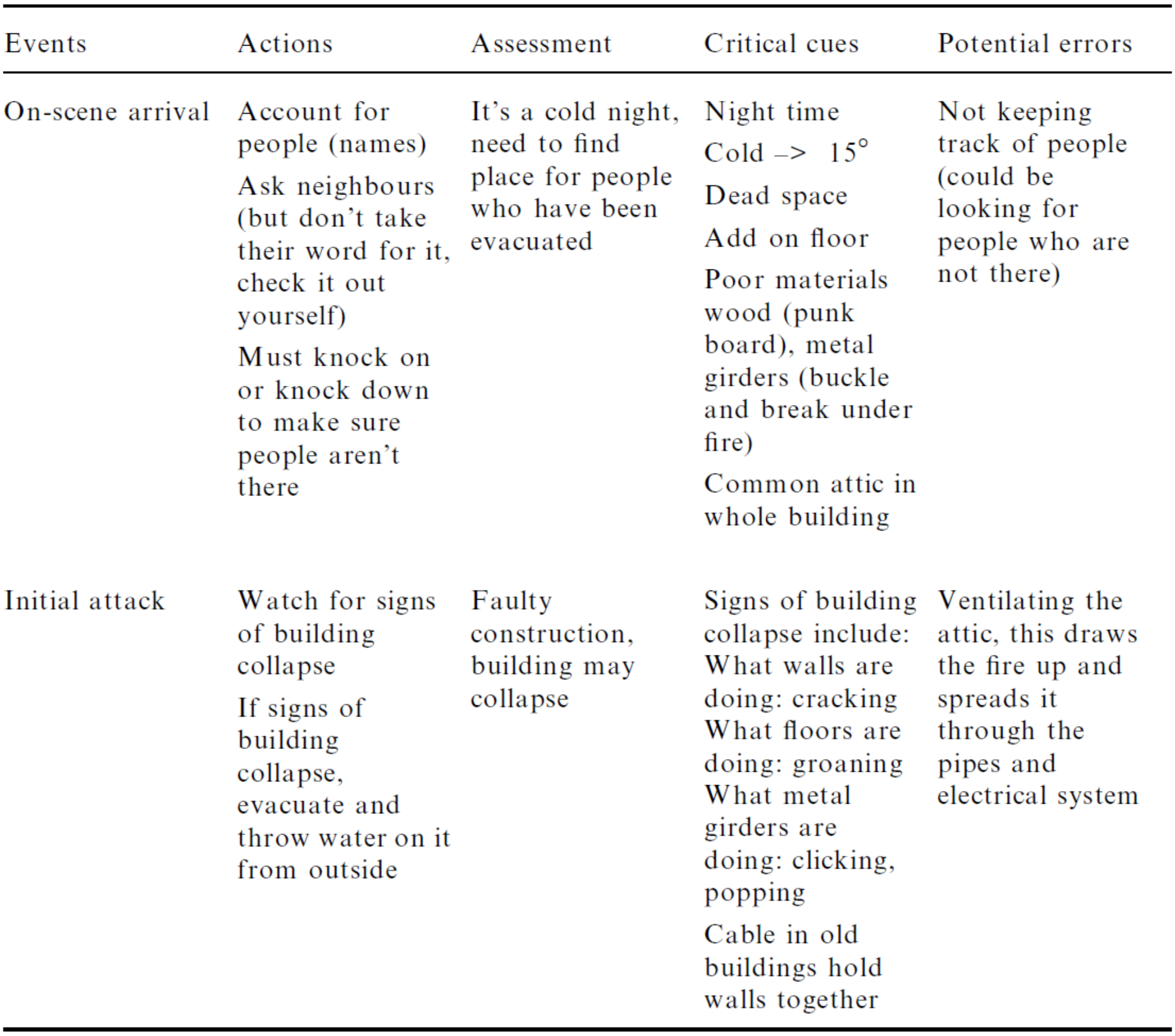
3.3.5 Managing qualitative data.

CDM interviews generate copious unstructured notes. Although there is a rich, scientific tradition for analyzing field notes and other qualitative data (Strauss & Corbin, 1994; Coffey et al., 1996), the Navy required efficient, pragmatic strategies for identifying aspects of expertise that could be integrated into curricula. ACTA addressed this challenge.

The note-taking forms for the Knowledge Audit (Figure 4) and Simulation Interview (Figure 5) were designed to streamline the analysis process. These matrices summarize aspects of expertise, illustrative examples, critical cues, actions, complexities, and common errors identified in each interview. The matrices facilitate data exploration to identify themes and idiosyncrasies across interviews without extensive data management and coding. During the interviews, Iencourage the use of a dedicated note-taker to record details that do not fit neatly into the note-taking matrices. These detailed notes and/or transcripts of the interview can be used to fill in details and context not captured in the matrices for more in-depth analysis.

3.4 ACTA Evaluation Study

To determine whether the streamlined ACTA methods could be used to effectively elicit aspects of expertise, I conducted a study in which I recruited volunteers naïve to CTA from psychology graduate programs to learn and administer the methods. Twelve students interviewed civilian firefighters and 11 interviewed Naval electronic warfare technicians. Subject matter experts and cognitive psychologists rated the materials generated by students to assess two aspects of validity: 1) did the materials include information that was predominantly cognitive in nature, and 2) was the information domain-specific and relevant? Raters were blind to the students' interview group. Findings from the evaluation study suggest that after a 6-hour workshop introducing the ACTA techniques, graduate students were able to elicit relevant, domain-specific, accurate cognitive information that was translated into training materials. (Militello & Hutton, 1998.)



Events	Actions	Assessment	Critical cues	Potential errors
On-scene arrival	Account for people (names) Ask neighbours (but don't take their word for it, check it out yourself) Must knock on or knock down to make sure people aren't there	It's a cold night, need to find place for people who have been evacuated	Night time Cold –> 15° Dead space Add on floor Poor materials wood (punk board), metal girders (buckle and break under fire) Common attic in whole building	Not keeping track of people (could be looking for people who are not there)
Initial attack	Watch for signs of building collapse If signs of building collapse, evacuate and throw water on it from outside	Faulty construction, building may collapse	Signs of building collapse include: What walls are doing: cracking What floors are doing: groaning What metal girders are doing: clicking, popping Cable in old buildings hold walls together	Ventilating the attic, this draws the fire up and spreads it through the pipes and electrical system

Figure 3.5: Sample Simulation Interview notetaking form (Militello & Hutton, 1998, p. 1624)

3.5 ACTA Epilogue

Since the ACTA suite of methods was published in 1998, it has been cited over 900 times according to Google Scholar. Appendix 3.1 lists 27 refereed journal articles that report using ACTA between 2000 and 2022. ACTA has been used to study 11 diverse domains expanding well beyond my body of work. Others have applied ACTA to study prescribing practices of nurse practitioners (Martini et al., 2022), search and rescue strategies of unmanned aircraft system pilots (Larcel & Andrews, 2021), use of an imaging system to maintain a picture of the external environment by submarine watch standers (Papautsky et al., 2020), global leadership (Brown et al., 2024), and decision strategies of rugby players (Johnston & Morrison, 2016).

It is also important to note that the ACTA methods have been extended in ways that I did not initially imagine. For example, the task diagram has been used to compare how different practitioners conceptualize their work. In a study of how primary care clinicians monitor and counsel patients regarding colorectal cancer (CRC) screening, researchers used the task diagram to highlight variations in the process across diverse health systems (Borders et al., 2014).

In training people to use the Knowledge Audit, I encourage them to tailor the probes to the domain they are studying. In this spirit, colleagues at Mitre developed a Human-Machine Teaming (HMT) Knowledge audit (McDermott et al., 2018), integrating knowledge audit probes with human-machine teaming requirements. It has since been used to study HMT in the context of space domain awareness (Fitzgerald et al., 2024). Others have combined the Knowledge Audit with the CDM to create the Critical Decision Audit (Borders & Klein, 2017).

These examples of ACTA applications and adaptations documented in the scientific literature suggest that the methods have had a practical role in uncovering new knowledge across a range of domains, and have served as an inspiration to researchers in tailoring methods to investigate DM in real-world contexts.

3.6 Beyond ACTA

Leading the effort to develop the ACTA suite of methods is my most cited contribution to the scientific community. Other important contributions include working with colleagues to develop a standard way of describing the CDM so that it could be trained to people who were not part of the apprenticeship program at Klein Associates. In addition to book chapters and articles (e.g., Jonassen et al., 1998; Klein et al., 2017), I have developed and facilitated CTA workshops for professional societies such as the Human Factors and Ergonomics Society, the American Medical Informatics Association, and for research organizations. I coauthored a leading textbook on CTA (see Hoffman & Militello, 2008). I developed and taught a semester-long course for graduate students at the University of Dayton.

In addition to training others to use CTA methods, I have been actively involved in efforts to adapt, expand, and combine CTA with other complementary methods such as ethnographic observation (Saleem et al., 2011), video analysis (Militello et al., 2018), cognitive work analysis (Militello et al., 2019a), and discrete event modeling (Militello et al., 2019b). I have also explored strategies for adapting CTA methods for use in large samples studies (Militello, Salwei, et al., 2023).

3.7 Summary and Conclusions

The ACTA suite of methods leveraged the strengths of the already established CDM to developed streamlined CTA methods. The resulting Task Diagram, Knowledge Audit, and Simulation Interview are a complementary set of methods. The ACTA methods have since been used in combination and individually to address a broad range of research questions. In addition to the continued application and adaptation of the ACTA methods, I have been actively involved in developing and conducting training for ACTA, CDM and other related. The next two chapters detail applications of CTA in training (Chapter 4) and technology design (Chapter 5) applications.

CHAPTER 4. COGNITVE TASK ANALYSIS TO INFORM TRAINING

4.1 Introduction

This chapter focuses on the use of CTA data to inform the design of training in high stakes domains. This type of training requires techniques that support the acquisition of recognition skills, or the ability to quickly size up a situation and know how to act (Militello, Sushereba, Ramachandran, 2023).

The chapter begins with a case study, *training pediatricians to recognize sepsis* (Section 4.2), illustrating three key points. First, it demonstrates the power of CTA for identifying expert cues and strategies to inform training. Second, this case study includes two types of efficacy data: a rigorous validation study and pragmatic use-based feedback. In combination, these suggest that the CTA-based sepsis training supports physicians in learning to recognize sepsis. Third, the case study demonstrates that our efforts to develop CTA workshops and how-to publications described in Section 3.5 have led to effective dissemination of the methods. Section 4.3 shifts focus to *training design principles* derived from an effort I led to integrate empirical and theoretical principles from disparate literatures. The resulting handbook offers guidance for designing training in high stakes domains (Militello, Sushereba, Ramachandran, 2023). Section 4.4 discusses two recent efforts to develop strategies for *assessing transfer a training* to on-the-job performance, a well-documented challenge (Kirkpatrick & Kirkpatrick 2016). The chapter closes with summary and conclusions.

4.2 Training pediatricians to recognize sepsis

I worked with a team of clinicians at Cincinnati Children's Hospital Medical Center (CCHMC) on a project funded by the Agency for Health Care Research and Quality to develop simulator-based training to support physicians in recognizing sepsis (Patterson et al., 2016; Geis et al., 2018). Sepsis is a systemic response to infection that can result in organ dysfunctions; it is a leading cause of death worldwide (Angus & Wax, 2001). Sepsis can be hard to diagnose because it often presents with subtle, nonspecific symptoms. Delayed recognition and management of sepsis is associated with increased morbidity and mortality (Han et al., 2003; Lundberg et al., 1998; Odetola et al., 2008). For this project, I trained a team of clinicians and behavioral researchers to conduct CDM, coached them by reviewing transcripts and offering suggestions for future interview sessions, and led the data analysis effort.

4.2.1 Methods

Data Collection

The team conducted 14 CDM interviews with four novices (interns), four senior trainees (senior residents), and six faculty (expert) physicians. Interviewees were asked to recall a patient for whom sepsis was suspected or later discovered. Interviews were recorded and transcribed. Figure 6 provides a sample timeline depiction of an incident created during a CDM interview.

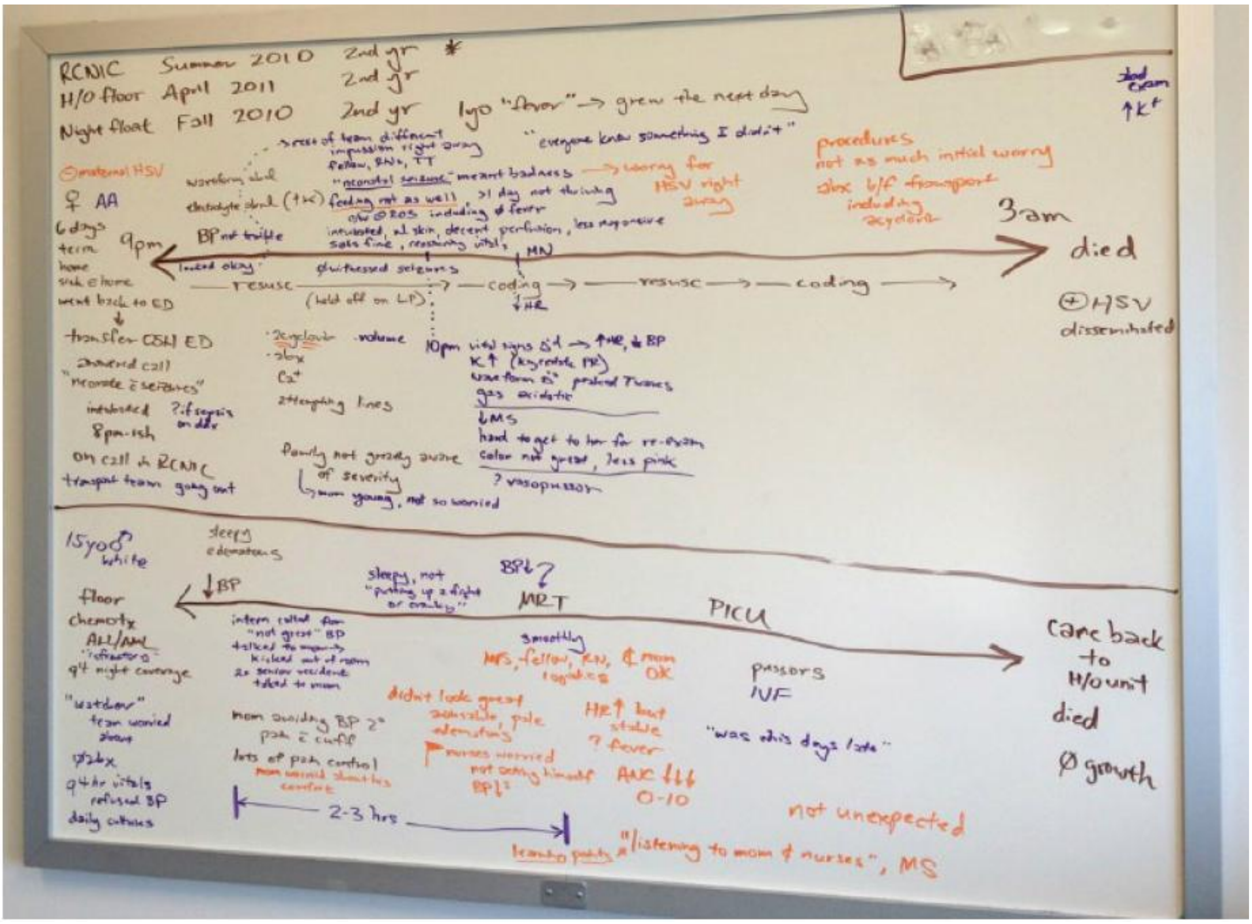
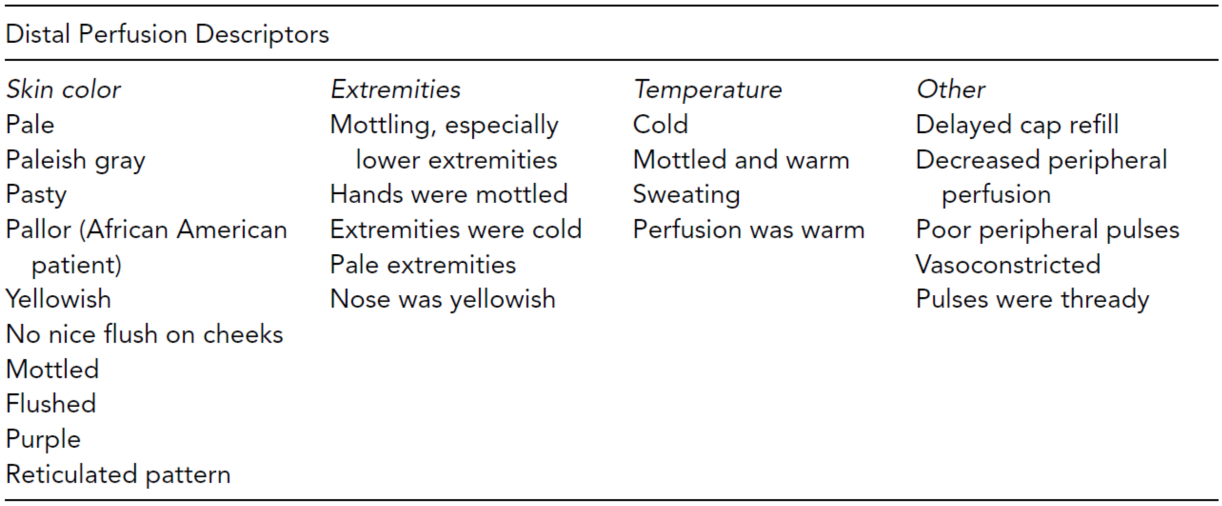


Figure 4.1: Incident timeline constructed during a CDM interview (Patterson et al., 2016, p. 41)

Data Analysis

The 14 interviews yielded 23 real-world incident accounts. Incident accounts were analyzed for cue recognition. At least 2 individual analysts reviewed each incident and extracted cues that were described by the interviewee. They then met to reach consensus. The result was a comprehensive list of cues noted in each case. Because capillary refill was identified as a particularly important cue that required judgment and was assessed in different ways, the analysis team created a critical cue inventory highlighting the range of cues described in assessing capillary refill (Table 4.1).

Table 4.1: Critical cue inventory detailing cues used to assess distal perfusion (Patterson et al., 2016), p. 47



Distal Perfusion Descriptors			
Skin color	*Extremities*	*Temperature*	*Other*
Pale	Mottling, especially lower extremities	Cold	Delayed cap refill
Paleish gray	Hands were mottled	Mottled and warm	Decreased peripheral perfusion
Pasty	Extremities were cold	Sweating	Poor peripheral pulses
Pallor (African American patient)	Pale extremities	Perfusion was warm	Vasoconstricted
Yellowish	Nose was yellowish		Pulses were thready
No nice flush on cheeks			
Mottled			
Flushed			
Purple			
Reticulated pattern			

Training Development

Using the incident accounts from the CDM interviews, the team worked with the simulation training department to develop six complementary training scenarios. To achieve cognitive fidelity, every effort was made to ensure that the cognitive challenges presented in the training scenarios were representative of challenges that would be encountered in the real world. The scenarios were built into the hospital's simulation facility. Simulation rooms were set up to replicate the emergency department. The team incorporated cues and clue clusters frequently described in the CDM interviews. Simulating some of these cues required the team to explore innovative strategies such as:

- Wrapping the hands and feet of the mannequin in ice prior to training to simulate cold extremities, an important cue indicating poor distal perfusion (see Table 4.1)
- Using a voice modulator so that training facilitators could speak through the mannequin in the voice of an infant, a toddler, or school-aged child, depending on the scenario to improve the fidelity of mental status cues
- Using videos and screen shots to depict changes in skin appearance and capillary refill rates; requiring the physician to interpret the image and draw his/her own conclusions.

Because several incident accounts included situations in which a patient's condition started at an early point in sepsis and slowly deteriorated, making it easier to miss subtle cues, the team designed scenarios with timelines more representative of early sepsis rather than the more acute phases typically included in training.

Training evaluation

We assessed training effectiveness in terms of validity measures, use, and perceived benefit. With regard to validity measures, the team conducted a study before the training scenarios were integrated into the resident training program. Eighteen physicians completed three sepsis simulations in the simulation facility plus a non-sepsis scenario (so that the correct answer would not always be sepsis). Performance of expert and novice physicians were compared.

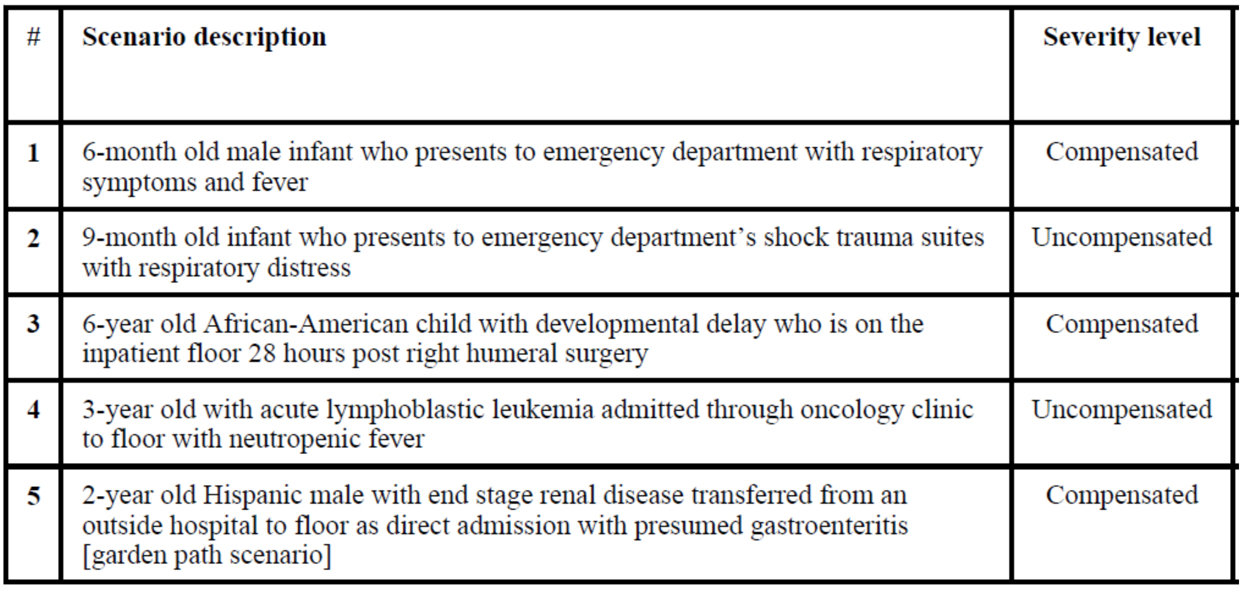
Use and perceived benefit were examined after the completion of the grant. With regard to use, CCHMC tracked use of the scenarios over 6 years to understand whether the organization continued to find them useful over time. With regard to perceived benefit, CCHMC asked faculty to provide feedback regarding whether residents became better at recognizing sepsis in the emergency department after participating in training.

4.2.2 Results

Training Scenarios

The six training scenarios combined sepsis-related cues in a variety of contexts. As part of the sepsis-focused scenarios, one "garden path" scenario was developed to present a seemingly straightforward case of hypovolemia, but as the scenario unfolds, additional information is presented suggesting sepsis. Table 4.2 summarizes the five sepsis scenarios. The sixth scenario included a child with supraventricular tachycardia (SVT), a condition that could easily be mistaken for sepsis.

Table 4.2: Summary of sepsis incidents (Geis et al., 2018), p. 19.



#	Scenario description	Severity level
1	6-month old male infant who presents to emergency department with respiratory symptoms and fever	Compensated
2	9-month old infant who presents to emergency department's shock trauma suites with respiratory distress	Uncompensated
3	6-year old African-American child with developmental delay who is on the inpatient floor 28 hours post right humeral surgery	Compensated
4	3-year old with acute lymphoblastic leukemia admitted through oncology clinic to floor with neutropenic fever	Uncompensated
5	2-year old Hispanic male with end stage renal disease transferred from an outside hospital to floor as direct admission with presumed gastroenteritis [garden path scenario]	Compensated

Evaluation results

Validation: A validation study provided evidence that experienced physicians were more likely to recognize sepsis in the simulated scenarios than residents (Geis et al., 2018). Sepsis was recognized in 19 (35%) of 54 simulations. The odds that experts recognized sepsis was 2.6 (95% CI 0.5, 13.8) times greater than novices. Adjusted for severity and level of expertise, the odds of recognizing sepsis was associated with an increase in the early sepsis management (ESM) checklist score of 1.8 (95% CI: 0.9, 3.6) and an increase in ESM global performance score of 4.1 (95% CI: 1.7, 10.0). These findings suggest that the scenarios provided practice at thinking like an expert in realistic, challenging situations.

Use: The sepsis scenarios became the foundation for Cincinnati Children's Hospital Medical Center's 3-year resident-focused, simulation-based training curriculum. The initial scenarios were expanded to include other conditions. They currently run the sepsis scenarios in 45 two-hour sessions each year. (G. Geis, personal communication, December 30, 2022)

Perceived benefit: Faculty reported that residents are better at recognizing sepsis due to this intervention. They also noted that residents do not cover patients in all units (surgical units, units covered only by APRNs, etc.). As a result, some units did not receive the sepsis training until recently when a series of sepsis cases were missed:

> *"We recently had a series of missed sepsis cases in these units, so the hospital has gone "full force" to focus on this [sepsis]. Last year we ran a mandatory course for RNs and APRNs on one of these units, which was so successful we are being tasked with training for all RNs and APRNs. We are leveraging our previous and current work with the*

> *residents, but have expanded the training to include VR and in situ sessions to complement training in the lab. It is all built around sepsis recognition.”*
>
> (G. Geis, personal communication, December 30, 2022)

4.2.3 Discussion and Conclusion

Although we cannot know for certain, this implementation of the sepsis training served as a natural experiment, suggesting that units where residents were not exposed to the sepsis scenarios were more likely to miss sepsis. The organization values these scenarios and believes that they do indeed prepare physicians to recognize this life-threatening condition.

This study highlights the value of CTA in eliciting real-world incidents that can be used as the basis for training scenario design with high cognitive fidelity. First, the CTA data provided insight into critical cues that experts rely on to recognize sepsis so they could be integrated into the training scenarios. Some of the cues identified in this study were not typically included in simulation-based training because they are not easily replicated in a simulation. As the study highlighted the role of these cues in expert assessment, the organization found innovative ways to include them in training. Second, the incident accounts served as the foundation for creating compelling real-world training scenarios. Third, interview data highlighted the importance of including more realistic timelines in the training so that learners could experience how sepsis evolves.

4.3 Emerging technology and recognition skills training

The sepsis study inspired me to explore how emerging technologies can be used to improve recognition skills training. I wanted to learn how augmented reality (AR) technology might be used to prepare learners to quickly size up a situation and know what to do. In 2016, I cofounded a company, Unveil, LLC, to develop augmented reality technology training for combat medics. We have developed training technology that is currently in use by civilian and military medical organizations, pairing AR with physical manikins in novel ways. Although no formal studies of efficacy have been conducted to date, testimonials suggest that instructors and students find the cognitive and perceptual fidelity offered by our training technology fills an important gap in medic training.

One criticism of AR-based training is that although it is rapidly proliferating, there is little evidence of its efficacy (Kaplan et al., 2020). This may have little to do with the potential training value of AR technology; rather, it more likely speaks to the limitations of poor training design. For those developing AR-based training, science-based guidance is difficult to find. Much of the theoretical and empirical research is published in edited volumes and conference proceedings. Journal articles are spread across the training, intelligent tutoring, NDM, psychology and other literatures. To meet the need for accessible science-based guidance, I worked with two coauthors to review disparate literatures and identify training design principles that are relevant to the design of recognition skills training. As lead author, I directed the literature review, led discussions of what to include and what to leave out, and served as the primary writer for the *Handbook of Augmented Reality Training Design* (Militello, Sushereba, Ramachandran, 2023). We combined the empirical and theoretical design principles with the experience of medic trainers, to articulate 11 training design principles (see Figure 4.2). Our intent is to help training designers move beyond the "wow factor" of AR and to leverage the strengths of AR for training recognition skills. The book has sold over 200 copies in the 18 months since publication, and received positive reviews from the medical simulation, human factors, and NDM communities (See Appendix 4.1 for book reviews).

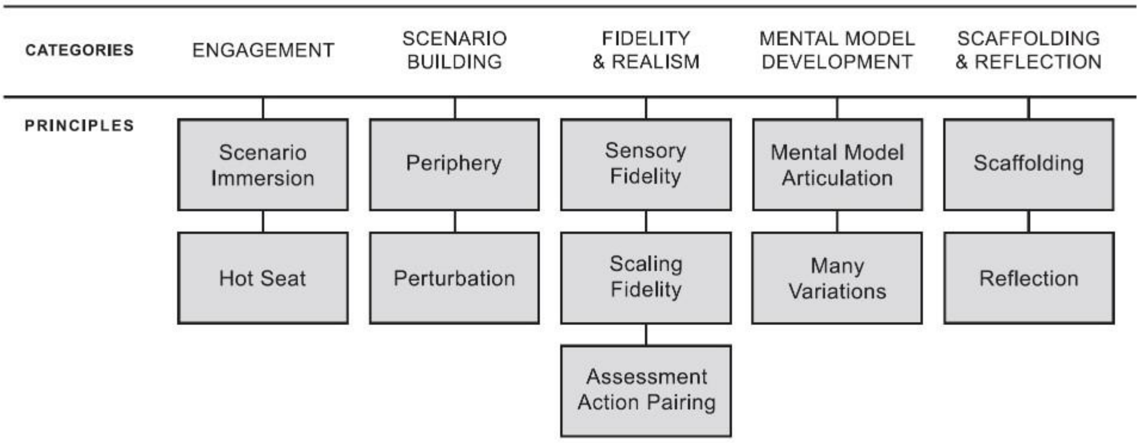


Figure 4.2: Eleven training design principles (Militello, Sushereba, Ramachandran 2023, p. 6)

4.4 Strategies for assessing transfer of training

Articulating empirical and theoretical design principles was an important step but finding strategies for effectively assessing how well training supports cognitive performance such as recognition skills was an equally important consideration. As cognitive performance in high stakes domains cannot be reduced to a single measure, effective evaluation is complex (Ernst et al., 2023; Patterson & Miller, 2010). Measures must be tailored to the cognitive demands identified in the CTA. I led two studies piloting measures to assess critical components of recognition skills training. The first study assessed a 30-minutes Tourniquet Troubleshoot Training (TTT) application (Sushereba, Militello, et al., 2024; Militello et al, 2023) and the second assessed a Training for Advanced Life Support in Austere Regions (TALSAR) (Sushereba, Fernandez, et al., 2024. Our approach included measuring four constructs: knowledge, recognition of errors/non-errors, transfer to novel situations, and transfer from mental practice to physical performance. Table 4.3 depicts examples of each measure as instantiated in TTT app evaluation (Sushereba, Militello. et al., 2024)

These pilot studies led to important lessons learned. They have helped us demonstrate that our measures of knowledge, recognition of error/non-errors, and transfer to novel situation measures are sensitive enough to differentiate between novice and skilled performers. This combination of measure provides a more nuanced understanding of training efficacy and limitations than the simple knowledge tests commonly used.

Measures of transfer from mental rehearsal to physical performance have been the most challenging to design. Although changes in actual on-the-job performance would be the gold standard of evaluation, I have no access to participants after the completion of our studies. Therefore, the focus has been on measures that require learners to demonstrate that they can transfer what they learned via mental rehearsal during training to performance in a simulated setting. To meet pragmatic constraints, our focus has been on rapid tests that are easily scored. For the TTT app study, this consisted of manikins with tourniquets already applied; learners were asked to assess the tourniquet and make any changes needed to correct errors. This measure was ineffective for several reasons: 1) there were only 2 tourniquets to be assessed so we had few data points, 2) the error was quite conspicuous, leading to a ceiling effect (nearly all control and

Table 4.3: Measures of task transfer

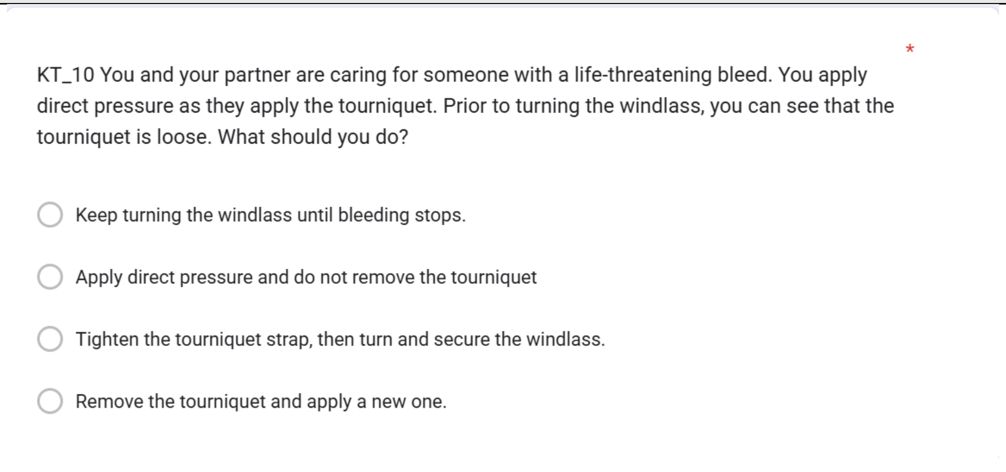
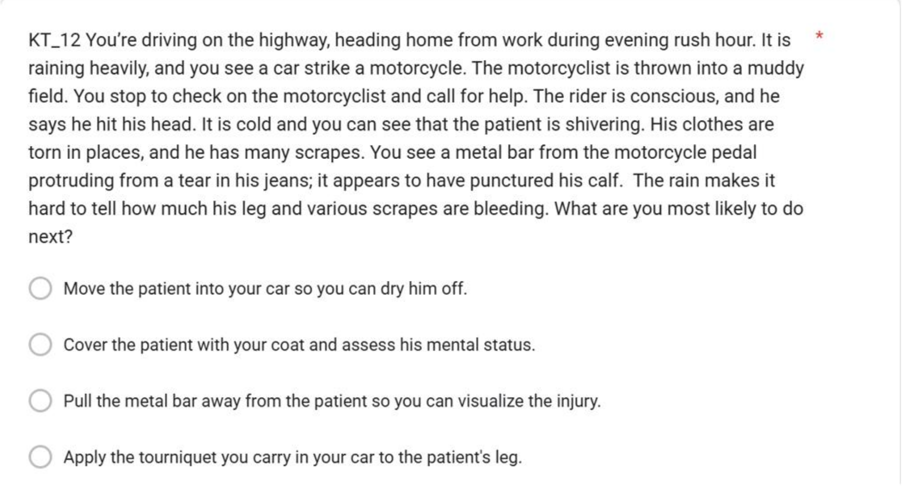
Construct	Measure	Example
Knowledge	Multiple choice knowledge test	 KT_10 You and your partner are caring for someone with a life-threatening bleed. You apply direct pressure as they apply the tourniquet. Prior to turning the windlass, you can see that the tourniquet is loose. What should you do? * ○ Keep turning the windlass until bleeding stops. ○ Apply direct pressure and do not remove the tourniquet ○ Tighten the tourniquet strap, then turn and secure the windlass. ○ Remove the tourniquet and apply a new one.
Recognition of errors/non-errors	Flashcard test of ability to distinguish errors from non-errors	 You come across someone with a leg injury that has been treated with a tourniquet. Do you see any problems with this tourniquet application? This injury does not require a tourniquet. There is no problem. The tourniquet isn't tight enough. The tourniquet is in the wrong location.

Table 4.3: Measures of task transfer continued

Construct	**Measure**	**Example**
Transfer to novel situations	Snapshot scenarios	 KT_12 You're driving on the highway, heading home from work during evening rush hour. It is raining heavily, and you see a car strike a motorcycle. The motorcyclist is thrown into a muddy field. You stop to check on the motorcyclist and call for help. The rider is conscious, and he says he hit his head. It is cold and you can see that the patient is shivering. His clothes are torn in places, and he has many scrapes. You see a metal bar from the motorcycle pedal protruding from a tear in his jeans; it appears to have punctured his calf. The rain makes it hard to tell how much his leg and various scrapes are bleeding. What are you most likely to do next? * ○ Move the patient into your car so you can dry him off. ○ Cover the patient with your coat and assess his mental status. ○ Pull the metal bar away from the patient so you can visualize the injury. ○ Apply the tourniquet you carry in your car to the patient's leg.
Transfer from mental practice to physical performance	Demonstration in simulator facility	

experimental subjects performed correctly), 3) the manikins did not bleed so some of the key cues were missing. We are currently engaged in an evaluation study, and have negotiated additional resources to carefully validate the physical demonstration measure prior to data collection.

4.5 Summary and Conclusions

CTA is a powerful tool for understanding the cognitive challenges of a domain to inform training and evaluation strategies, particularly for recognition skills in high stakes domains (Militello, Sushereba, Ramachandran., 2023). The real-world application of CTA to develop sepsis recognition training is a compelling example of the use of CTA to support complex skill acquisition. The potential benefits of CTA to inform training may be amplified by AR technology that provides a platform for presenting learners immersive, photo-realistic cues in a portable form. The *Handbook of Augmented Reality Training Design* (Militello et al., 2023) offers empirical and theory-based training design principles to support training developers in creating effective recognition skills training. Assessing transfer of training to on-the-job performance is notoriously difficult with recognition skills because much of the cognitive skill is invisible to observation. I have led two pilot study that have successfully demonstrated the power of combining measures of knowledge, recognition of errors/non-errors, and transfer to novel situations to assess the efficacy and limitations of training. Effective and practical measures of transfer from mental practice to physical performance are currently under development.

CHAPTER 5. CTA TO INFORM TECHNOLOGY DESIGN

5.1 Introduction

One common criticism of technology is that it is designed to support “work as imagined” rather than “work as done” (Hollnagel & Woods, 1983; Miller & Feigh, 2019; Shorrock, 2022; Woods & Dekker, 2000). There is a tendency to design based on use cases that describe typical operations. However, in highly dynamic work settings, typical operations are few and far between. Because CTA is focused on understanding expertise in context, including during non-routine events, it can provide important guidance to technology design.

This chapter includes a description of the Decision-Centered Design (DCD) framework (Section 5.2), articulated to guide researchers and technology developers in applying what they learn from CTA to design (Militello & Klein, 2013). DCD is distinct from other design frameworks in that it emphasizes designing to support tough decisions, rather than focusing primarily on routine operations. The DCD framework serves as an important tool to communicate the contributions of CTA to multi-disciplinary design team. Section 5.3 describes a case study employing CTA to inform the design and evaluation of a decision support application for CRC screening (Militello et al., 2016; Militello et al., 2017). This case study illustrates a design strategy that was novel at the time: mining the data in the electronic health record (EHR) and presenting it in succinct representations that support DM. Section 5.4 briefly summarizes the Integrated Cognitive Analysis (ICA) framework that incorporates CTA with other cognitive engineering frameworks to support the design of first-of-kind systems that include advanced automation and autonomy (Ernst et al., 2019, in press; Militello et al., 2019b).

5.2 Decision-Centered Design

As CTA practitioners began to get a voice at the table in multi-disciplinary design and development teams, it became clear that we needed to articulate a framework that would help others understand the contributions of CTA. The DCD framework highlights three contributions of CTA to support effective collaboration in multi-disciplinary teams. First, DCD emphasizes the use of CTA methods to uncover expertise and decision requirements, elements of work that are commonly overlooked. Second, DCD articulates how CTA findings play a key role in the design of technology that supports human-machine teaming. Third, DCD explains how CTA findings inform evaluation strategies that assess how well a proposed design intervention supports cognitive activities (Crandall et al., 2006; Hutton et al., 2003; Militello & Klein, 2013).

As we developed workshops to train others to conduct CTA (see Chapter 3), we reflected on the steps involved in our projects focused on technology design. Figure 5.1 summarizes the five stages.

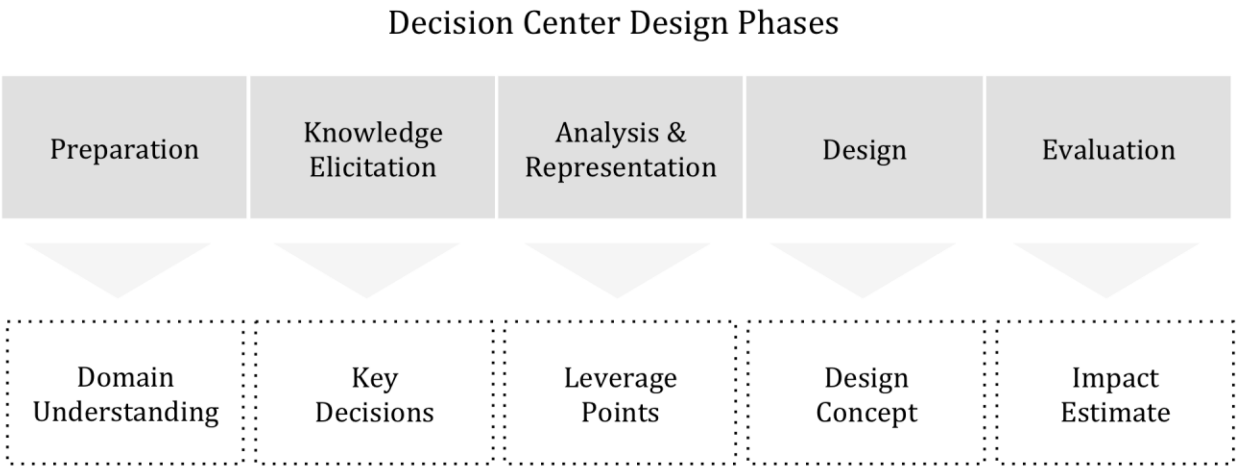


Figure 5.1: Overview of Decision-Centered Design phases. Figure adapted from Crandall et al., 2006, p. 181

DCD can be defined in this way:

> “Decision-Centered Design advocates for designs that focus on difficult decisions and unexpected situations rather than routine operations. Decision-Centered Design focuses on identifying key decisions rather than exhaustively documenting all possible cognitive requirements…. Decision-Centered Design encourages incident-based evaluation strategies that are context specific...” (Militello & Klein, 2013, p.261)

The DCD framework has been used to support multi-disciplinary design teams in a range of contexts including damage control system design for Navy ships (Miller et al., 2003), clinical decision support (Harle, et al., 2019; Assadi et al., 2022), design implications for automation in petrochemical plant control rooms (Wang et al., 2024), and user interfaces for the electric power industry (Gualtieri et al., 2012).

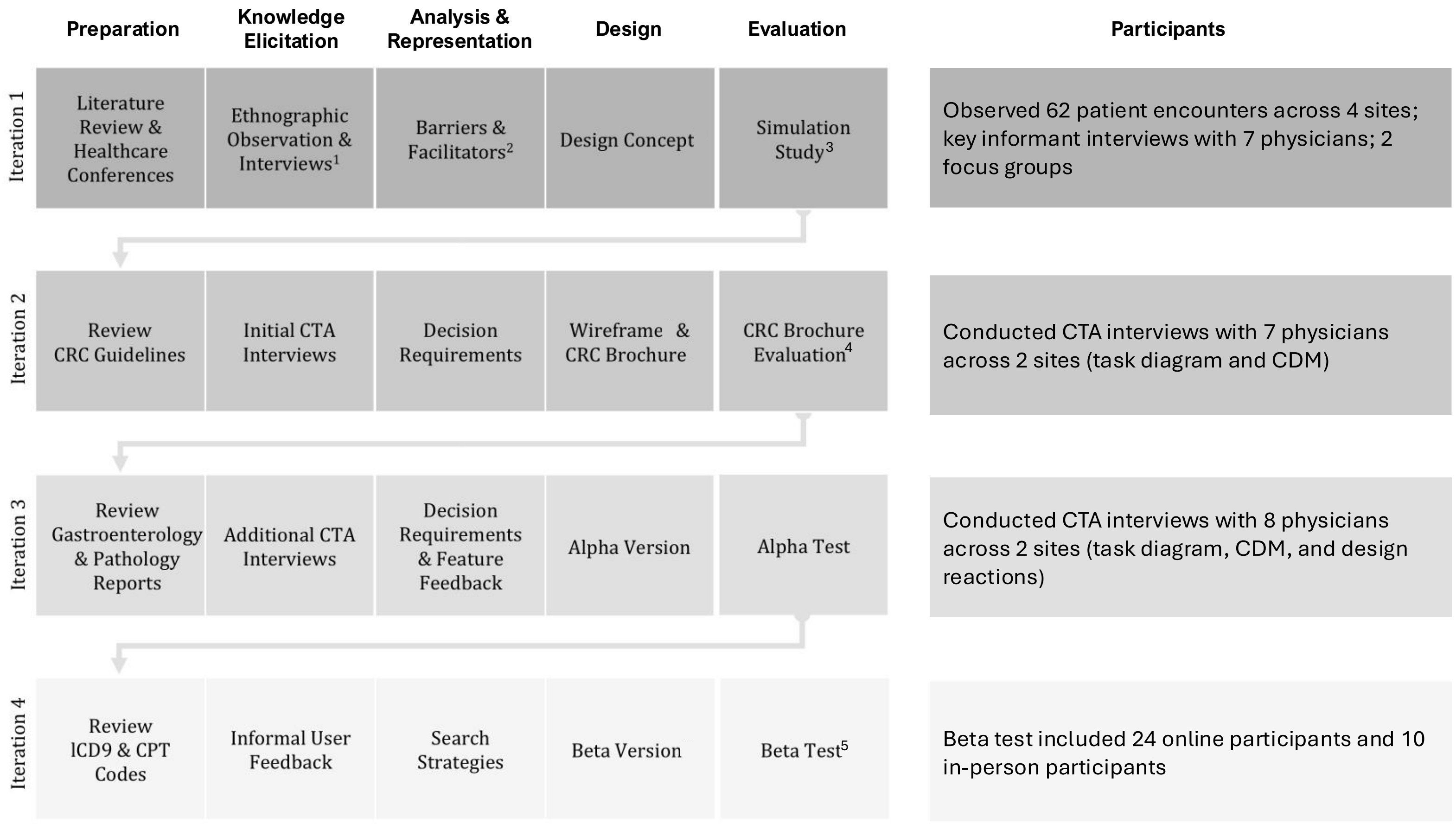
5.3 Designing Decision Support for Colorectal Cancer Screening: A Case Study

I led a team to develop decision support for primary care clinicians in managing CRC screening for their patients (funded by U.S. Centers for Disease Control & Prevention, 2010-2014). This work occurred during a period of rapid adoption of electronic health records in the U.S. Although the potential for integrating decision support into EHRs was widely touted, efforts had met with limited success. Many highlighted a misalignment between the types of decision support features offered, the cognitive processes of health care professionals, and the work environment (Sidebottom et al, 2012; Streiff et al., 2012; Tawfik et al, 2012).

At the time of this study, CRC screening rates remained stubbornly low despite the introduction of clinical reminders. In 2012, only 65% of eligible adults in the U.S. were up-to-date with CRC screening; 7% had been screened but were not up-to-date, and 28% had never been screened (Centers for Disease Control and Prevention, 2013).

5.3.1 CTA to identify cognitive requirements

I led an iterative CTA, including observations of clinicians counseling patients about CRC screening, and interviews with primary care clinicians across a range of health settings. Each data collection effort deepened and broadened our understanding of the cognitive challenges primary care clinicians faced. Figure 5.2 summarizes each iteration.



	Preparation	Knowledge Elicitation	Analysis & Representation	Design	Evaluation	Participants
Iteration 1	Literature Review & Healthcare Conferences	Ethnographic Observation & Interviews[1]	Barriers & Facilitators[2]	Design Concept	Simulation Study[3]	Observed 62 patient encounters across 4 sites; key informant interviews with 7 physicians; 2 focus groups
Iteration 2	Review CRC Guidelines	Initial CTA Interviews	Decision Requirements	Wireframe & CRC Brochure	CRC Brochure Evaluation[4]	Conducted CTA interviews with 7 physicians across 2 sites (task diagram and CDM)
Iteration 3	Review Gastroenterology & Pathology Reports	Additional CTA Interviews	Decision Requirements & Feature Feedback	Alpha Version	Alpha Test	Conducted CTA interviews with 8 physicians across 2 sites (task diagram, CDM, and design reactions)
Iteration 4	Review ICD9 & CPT Codes	Informal User Feedback	Search Strategies	Beta Version	Beta Test[5]	Beta test included 24 online participants and 10 in-person participants

Figure 5.2: CTA instantiated in an iterative DCD process. Adapted from Militello, et al., 2016, p. 76. Numbered superscript annotations refer to publications describing individual iterations: [1]Saleem et al. (2009). [2]Saleem et al. (2005). [3]Saleem et al. (2011). [4]Borders et al. (2014). [5]Militello et al. (2017).

Iteration 2 CTA (see Figure 5.2) identified six key cognitive requirements to drive our design, listed in Table 5.1.

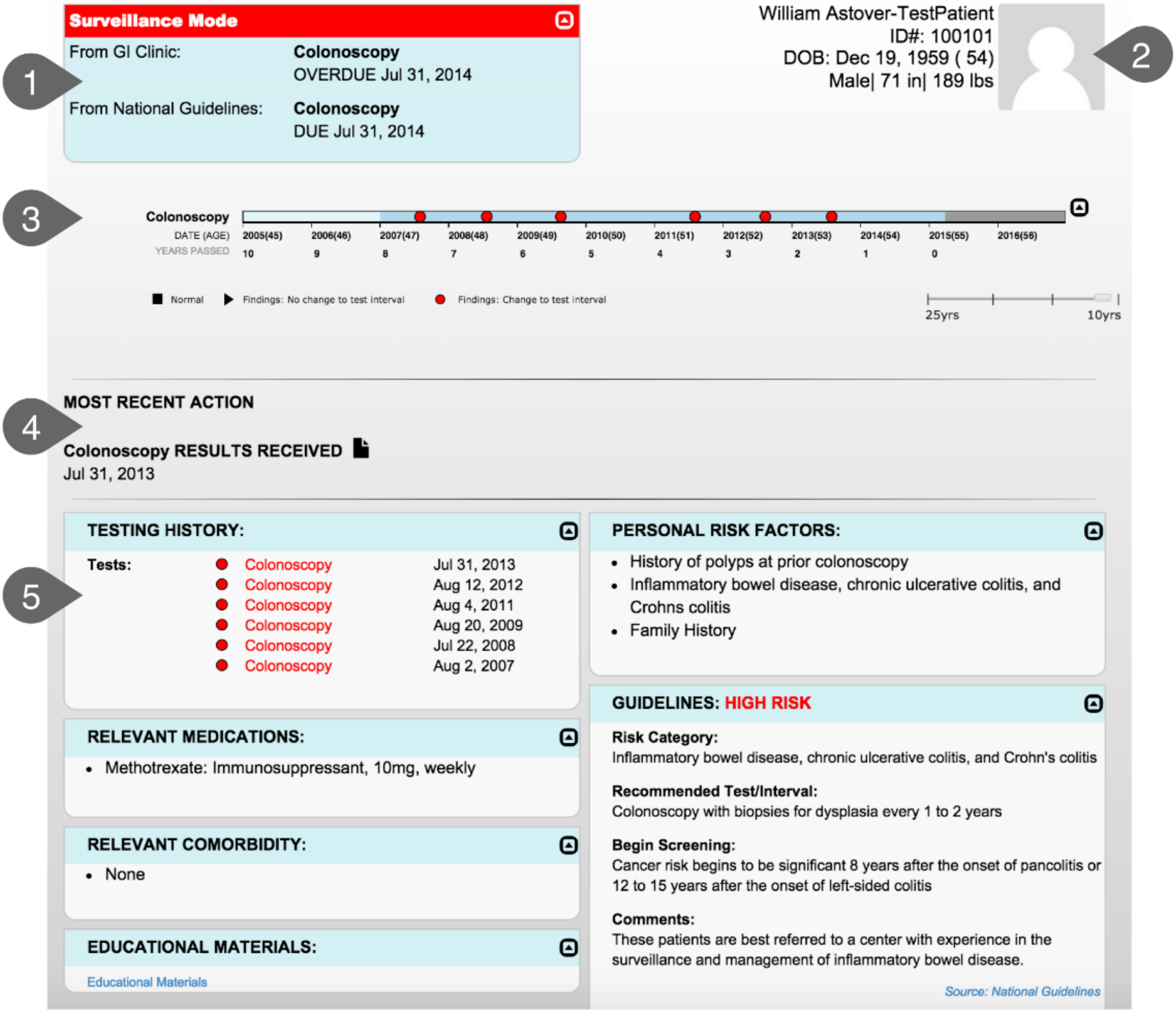
Table 5.1: Key cognitive requirements for managing CRC screening

Cognitive support requirement	**Description**
1. Determine whether the patient is in screening or surveillance mode	For experienced primary care providers, this is an important sensemaking frame. Screening versus surveillance mode has important implications for what information the provider will access prior to the patient discussion as well as how the provider will present and discuss testing options with the patient. For those in surveillance mode, it is important to review prior findings and to ensure that the patient understands that prior findings could increase the patient's risk for CRC. Thus, the importance of further testing at recommended intervals is greater. For those in screening mode, no additional information gathering is generally needed, and the conversation may be simpler.
2. Obtain a big picture perspective of the patient's testing history	Prior CRC test data may be found in progress notes, lab reports, gastroenterology reports, and pathology reports. During a patient encounter, it is difficult to use the EHR to access each of the required screens, locate the relevant information, and mentally integrate the data in a timely manner while talking to the patient. In fact, in some cases finding and integrating relevant data is enough of a barrier that physicians rely on patient memory of past CRC tests and findings rather than search the EHR. This represents a considerable barrier to effective sensemaking.
3. Know where the patient is in the screening cycle	In most cases, the primary care provider orders a test and receives a report from the specialty clinic or lab in a few weeks. In some cases, however, the primary care provider does not receive a report. In these cases, it is difficult to determine where the process fell apart. There is generally no visibility into what happens after the test is ordered and why a test did not occur, greatly hindering the provider's ability to detect problems with the process.
4. Consider conditions or medications that have implications for CRC testing	Quickly reviewing relevant conditions and medications helps the primary care provider recognize non-routine situations and make patient-based recommendations for testing. For example, some primary care providers reported that they consider whether the patient has a condition that may increase the risk associated with the anesthesia often used with colonoscopy. For those patients, they may recommend another test modality.

Table 5.1: Key cognitive requirements for managing CRC screening continued

	increase the risk associated with the anesthesia often used with colonoscopy. For those patients, they may recommend another test modality.
5. Assess and monitor a patient's individual risk level for CRC	Information related to risk stratification might be found in multiple places in the EHR, including prior progress notes and GI reports. Furthermore, risk stratification may change based on test findings or even changes in family history (e.g., a first-degree relative recently diagnosed with CRC). Primary care providers indicated that it would be useful to have the most recent data relevant to risk level available so they can quickly assess and ask relevant questions to determine whether there is a need to update their understanding of the patient's CRC risk.
6. Educate and inform patients	Primary care providers report that they want each patient to understand what colon cancer is, and what the screening options are. Most report that they emphasize colonoscopy as a gold standard of care because it provides a more complete view of the colon and because the gastroenterologist is able to remove precancerous polyps during the procedure. Primary care providers report common CRC misconceptions from patients include underestimation of the risk of colorectal cancer, overestimation of the risk of colonoscopy procedure, fear that the colonoscopy procedure will be uncomfortable, and belief that CRC screening is expensive.

CTA with primary care clinicians revealed that much of the information needed to support cognitive requirements 1-5 was available in the EHR, but difficult to find. Rather than spend precious time searching the EHR, clinicians would frequently ask patients (often unreliable reporters) about their screening history. Our solution was to design software that would locate data in the EHR and display it in an at-a-glance visualization containing key information related to cognitive requirements 1-5, with direct links to more detailed reports and resources (Figure 5.3). We also designed patient-facing educational materials to support cognitive requirement 6 (Figure 5.4) (Militello et al., 2016).



Surveillance Mode

From GI Clinic:	**Colonoscopy** OVERDUE Jul 31, 2014
From National Guidelines:	**Colonoscopy** DUE Jul 31, 2014

William Astover-TestPatient
ID#: 100101
DOB: Dec 19, 1959 (54)
Male| 71 in| 189 lbs

MOST RECENT ACTION

Colonoscopy RESULTS RECEIVED
Jul 31, 2013

TESTING HISTORY:

Tests:		
	● Colonoscopy	Jul 31, 2013
	● Colonoscopy	Aug 12, 2012
	● Colonoscopy	Aug 4, 2011
	● Colonoscopy	Aug 20, 2009
	● Colonoscopy	Jul 22, 2008
	● Colonoscopy	Aug 2, 2007

RELEVANT MEDICATIONS:

- Methotrexate: Immunosuppressant, 10mg, weekly

RELEVANT COMORBIDITY:

- None

EDUCATIONAL MATERIALS:

Educational Materials

PERSONAL RISK FACTORS:

- History of polyps at prior colonoscopy
- Inflammatory bowel disease, chronic ulcerative colitis, and Crohns colitis
- Family History

GUIDELINES: HIGH RISK

Risk Category:
Inflammatory bowel disease, chronic ulcerative colitis, and Crohn's colitis

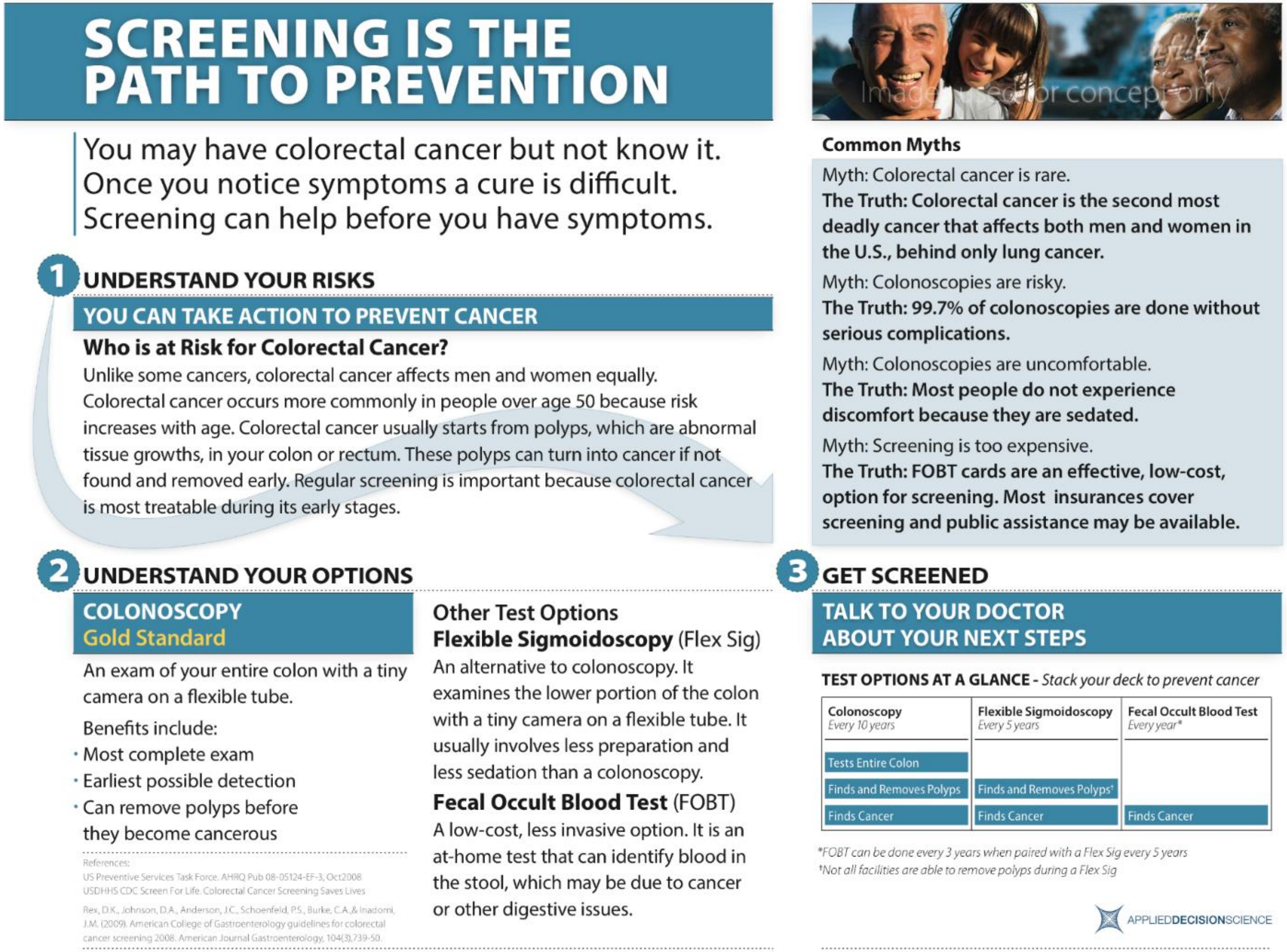
Recommended Test/Interval:
Colonoscopy with biopsies for dysplasia every 1 to 2 years

Begin Screening:
Cancer risk begins to be significant 8 years after the onset of pancolitis or 12 to 15 years after the onset of left-sided colitis

Comments:
These patients are best referred to a center with experience in the surveillance and management of inflammatory bowel disease.

Source: National Guidelines

Figure 5.3: Beta version of the Screening & Surveillance App. 1.0 = recommendations at a glance; 2.0 = patient demographics; 3.0 = timeline; 4.0 = most recent action panel; 5.0 = dashboard. (Militello et al., 2016, p. 83). Note: William Astover is a fictional test patient; no actual patient data are displayed in this figure.



SCREENING IS THE PATH TO PREVENTION

You may have colorectal cancer but not know it. Once you notice symptoms a cure is difficult. Screening can help before you have symptoms.

1 UNDERSTAND YOUR RISKS

YOU CAN TAKE ACTION TO PREVENT CANCER

Who is at Risk for Colorectal Cancer?

Unlike some cancers, colorectal cancer affects men and women equally. Colorectal cancer occurs more commonly in people over age 50 because risk increases with age. Colorectal cancer usually starts from polyps, which are abnormal tissue growths, in your colon or rectum. These polyps can turn into cancer if not found and removed early. Regular screening is important because colorectal cancer is most treatable during its early stages.

2 UNDERSTAND YOUR OPTIONS

COLONOSCOPY
Gold Standard

An exam of your entire colon with a tiny camera on a flexible tube.

Benefits include:

- Most complete exam
- Earliest possible detection
- Can remove polyps before they become cancerous

References:

US Preventive Services Task Force. AHRQ Pub 08-05124-EF-3, Oct2008

USDHHS CDC Screen For Life. Colorectal Cancer Screening Saves Lives

Rex, D.K., Johnson, D.A., Anderson, J.C., Schoenfeld, P.S., Burke, C.A.,& Inadomi, J.M. (2009). American College of Gastroenterology guidelines for colorectal cancer screening 2008. American Journal Gastroenterology, 104(3),739-50.

Other Test Options

Flexible Sigmoidoscopy (Flex Sig)

An alternative to colonoscopy. It examines the lower portion of the colon with a tiny camera on a flexible tube. It usually involves less preparation and less sedation than a colonoscopy.

Fecal Occult Blood Test (FOBT)

A low-cost, less invasive option. It is an at-home test that can identify blood in the stool, which may be due to cancer or other digestive issues.

Image used for concept only

Common Myths

Myth: Colorectal cancer is rare.
The Truth: Colorectal cancer is the second most deadly cancer that affects both men and women in the U.S., behind only lung cancer.

Myth: Colonoscopies are risky.
The Truth: 99.7% of colonoscopies are done without serious complications.

Myth: Colonoscopies are uncomfortable.
The Truth: Most people do not experience discomfort because they are sedated.

Myth: Screening is too expensive.
The Truth: FOBT cards are an effective, low-cost, option for screening. Most insurances cover screening and public assistance may be available.

3 GET SCREENED

TALK TO YOUR DOCTOR ABOUT YOUR NEXT STEPS

TEST OPTIONS AT A GLANCE - *Stack your deck to prevent cancer*

Colonoscopy *Every 10 years*	**Flexible Sigmoidoscopy** *Every 5 years*	**Fecal Occult Blood Test** *Every year**
Tests Entire Colon		
Finds and Removes Polyps	Finds and Removes Polyps†	
Finds Cancer	Finds Cancer	Finds Cancer

**FOBT can be done every 3 years when paired with a Flex Sig every 5 years*

†Not all facilities are able to remove polyps during a Flex Sig

APPLIEDDECISIONSCIENCE

Figure 5.4: Patient-facing educational materials (Militello et al., 2016, p.82)

5.3.2 Evaluating the Screening and Surveillance App

The team conducted a two-part study to evaluate the Screening & Surveillance Application (SSA) (Militello et al., 2017). Part 1 was an online study with a broad sample of primary care clinicians across the U.S, who were experienced with a range of EHRs (n=24). Part 2 was in-person study with primary care clinicians at a Veteran's Health Administration Medical Center (n=10). For both the online and in-person studies, participants worked through mock patient scenarios, and answered questions related to workload (online) or mental effort (in-person) and usability. Findings from these studies suggested that the SSA was effective in supporting clinicians in managing CRC screening for their patients. Participants were able to answer questions about CRC related patient data accurately using the SSA. The SSA required less workload than the EHR alone and participants completed patient scenarios 29% faster with the SSA. Participants rated the SSA highly usable and useful. Table 5.2 provides details from the in-person study comparing the SSA to the EHR.

5.3.3 Contributions of the Screening & Surveillance App

Our research project to design, develop, and evaluate the SSA for CRC screening provided three original contributions. First, it demonstrated the power of CTA and DCD for identifying cognitive requirements to inform and guide design. The second innovation was building a modular application that leveraged data already stored in the EHR, thereby making it more accessible. This innovative strategy is commonly used today. Third, our evaluation included both an online and an in-person component, allowing us to collect complementary data which provided a more comprehensive understanding of the strengths and limitations of the SSA.

5.4 Integrated Cognitive Analysis for Human-Machine Teaming

Many CTA projects focus on designing tools such as the SSA app to support relatively narrow tasks. In some cases, however, we are asked to apply CTA to the design of complex systems that represent revolutionary changes in work, extending the notion of DCD. For example, I led the optimally crew vehicle project (Militello et al., 2019a) funded by the U.S. Army to make recommendations regarding how many crew members and what types of automation would be needed for specific missions using future helicopters (Militello et al., 2019b). For these

Table 5.2: Findings from in-person study comparing SSA to EHR (Militello et al., 2017, p. 771).

Measure	**SSA Mean (SD)**	**EHR Mean (SD)**	***t*(df)**	***p***
Performance				
Accuracy	9.15 (0.78)	6.95 (1.19)	6.14 (9)	< 0.001
Time	187.31 (57.18)	262.90 (63.81)	-4.42 (9)	= 0.002
Screens Accessed	3.45 (0.76)	10.45 (3.49)	-6.36 (9)	< 0.001
Mouse Clicks	15.06 (4.89)	24.56 (10.98)	2.32 (7)	= 0.053
Mental Effort				
RSME Patient Scenarios	20.78 (8.28)	51.88 (20.72)	-4.42 (9)	= 0.002
RSME Retrospective	16.90 (8.22)	55.65 (19.36)	-7.77 (9)	< 0.001
Usability (5-point scales)				
Health ITUES Avg.	4.67 (0.37)	-	-	-
Quality of Work Life	4.67 (0.44)	-	-	-
Perceived Usefulness	4.59 (0.46)	-	-	-
Perceived Ease of Use	4.83 (0.31)	-	-	-

"envisioned world" problems, CTA can be integrated with other cognitive engineering methods. One early criticism of CTA methods was that because they tend to use retrospective interview techniques, they only contribute to our understanding of current operations. This project demonstrated how CTA can be used to identify existing cognitive challenges *and* highlight cognitive challenges in the envisioned world.

5.4.4 Designing for the envisioned world

The Integrated Cognitive Analysis (ICA) for human-machine teaming (Ernst et al., 2019; Ernst et al., in press) includes a set of methods for foundational analysis to understand the work and generate design concepts, and a set of evaluative analysis methods to iteratively evaluate and refine design concepts (Figure 5.5). This methodology integrates cognitive engineering methods that are rarely used in combination. For the challenges of designing for a future world in which humans are interacting with advanced automation and autonomy, the ICA leverages the strengths of the individual methods and amplifies them. The methods complement and strengthen each other in important ways. For example, one criticism of discrete event modeling approaches is that they tend to oversimplify phenomena by relying on readily available data. Using CTA to inform models leads to models that include more real-world complexities. Similarly, many find that work analysis representations have limited effectiveness as a communication tool, but when combined with vignettes, examples, and complexities elicited in CTA interviews, they become a useful tool for establishing common ground across the design team. Conversely, CTA practitioners often have difficulty linking interview data to implications for design at a system level; used in combination with work analysis, this problem is reduced.

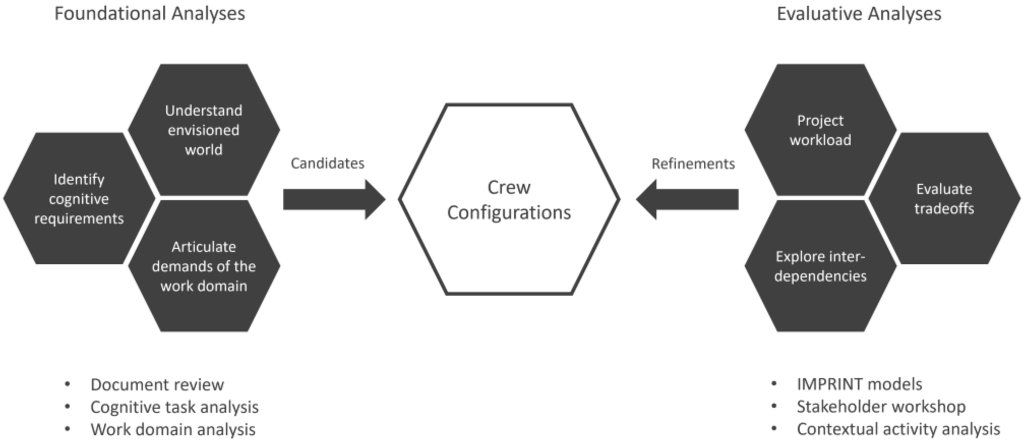


Figure 5.5: Integrated cognitive analysis for human-machine teaming (Ernst et al., in press, p. 3).

5.4.5 ICA-HMT Contributions

The difficulties of designing for the envisioned world include uncertainty about technology readiness, crewing, and the challenges human-machine teams will face. ICA combines cognitive engineering methods to support researchers in considering human-machine teaming in the early conceptual design phase. ICA has been used successfully in two contexts. For the OCV project, ICA led to recommendations cautioning Army leadership that full crewing would be required for complex, critical attack and reconnaissance missions (Militello et al., 2019b). Reduced crewing with greater reliance on cockpit automation would likely be feasible in the future for highly predictable missions such as resupply. ICA was later used by NASA to make crewing recommendations for Mars missions (Dempsey et al., 2024).

5.5 Summary and Conclusions

Integrating CTA into design often includes collaboration with software developers, mechanical engineers, program managers, and others. I have helped promulgate the DCD framework to aid other disciplines in understanding where CTA fits into the design process and the contributions that can be realized. The case study of creating decision support for CRC screening management illustrates how CTA served as a core component of each iteration of design, and led our team to a novel approach to supporting clinician DM. The ICA extends CTA beyond initial applications to understand difficult-to-articulate aspects of work, to aid experts in extrapolating from current operations to an envisioned future world. The ICA is an important step toward addressing the need to integrate CTA with other cognitive analyses to design for a world that includes smart automation and autonomy. The primary drawback of the ICA is that it is labor intensive. Future efforts will explore strategies for adapting ICA for use at scale as the need to support human-machine teaming in envisioned worlds is likely to continue to grow in the coming years. The next chapter provides a reflection on all three research themes presented in this synoptic report.

CHAPTER 6. SYNOPSIS AND NEXT STEPS

6.1 Introduction

This chapter provides a critical synopsis of the three CTA-related themes discussed in this synoptic report. It addresses next steps in the refinement and application of CTA with an emphasis on the role of CTA in designing for advanced automation and artificial intelligence. Please see Appendix 6.1 for a reflective account of my research journey.

6.2 Consolidation of Research

The three themes of this synoptic report were largely driven by societal challenges. The first theme, *design and development of CTA methods*, was largely driven by the fact that I began my career at a time when research was funded to develop methods needed to meet the needs of the 'information age'. Specifically, there was a call for methods that would drive the design of tools and technology needed to support human workers as more physical and predictable jobs were taken over by machines. The ACTA methods have had greater reach than ever anticipated. My efforts to disseminate ACTA and other CTA methods by publication, workshops, and online self-study has lent credibility to these methods as researchers around the world have applied and adapted them to answer a range of research questions and address applied problems.

The second theme, *applying findings from CTA to inform training*, was driven by changes in training practices, the accessibility of AR technology, and a need to better understand how training cognitive skills transfers beyond the training experience. I trained a team of pediatricians to conduct CTA and worked with them to analyze data to inform training for sepsis recognition; the resulting training is still in use 6 years after the study. This work was in direct response to changes in medical training that left resident physicians unprepared to recognize sepsis. As AR technology became readily available and affordable, I led an effort to abstract training design principles from NDM and related literatures. *The Handbook of Augmented Reality Training Design Principles* focuses on how to leverage the strengths of AR to provide recognition skills for high stakes domains. I led efforts to develop strategies for evaluating transfer of training to novel situation and from mental rehearsal to physical actions. Questions of transfer become increasingly important as more training takes place using AR and virtual reality.

The third theme, *applying findings from CTA to inform technology design*, was driven by the need to bridge the gap between CTA findings and design. As CTA methods gained traction, there was a still a need to help multidisciplinary teams understand CTA methods and how they can contribute to design and evaluation. Decision-centered design is a framework articulated to meet this need. As the U.S. government incentivized rapid adoption of EHRs, many medical professionals found themselves foraging in the health record, searching for information they needed to counsel patients about CRC screening. I led a project to using CTA to elicit information needed to make decisions about CRC screening. We used the DCD framework and CTA findings to design and evaluate an application that would extract screening history, comorbidities, and other relevant information from the EHR and display it in a clear, concise visualization. In 2019, the U.S. Army authorized the design and development of helicopters that would use advanced automation and autonomy to fly further and faster than any existing rotorcraft. They contacted me with a request to lead an effort to use CTA to identify the cognitive demands for pilots in this envisioned world and make recommendations about crewing configurations in different mission contexts. I assembled a team of cognitive engineers with complementary expertise. We articulated the ICA for human machine teaming, and exercised it over the course of a 9-month project. The ICA has since been adapted for use by NASA to explore crewing recommendations for Mars missions.

6.3 Original Contributions

Seven contributions to knowledge and practical impact are detailed in Table 1.1. In this section, I highlight the ones I perceive to be most significant. First, the ACTA methods represent an original scientific methodological contribution, and they have been used to support significant contributions to research worldwide. Appendix 3.1 lists a broad range of applied research projects that have used ACTA. Second, the sepsis recognition study (Patterson et al, 2016) demonstrates that the workshops we developed (Klein et al., 2017), can be used effectively to train others to use CTA. The successful collaboration between CTA practitioners (me) and domain experts (pediatricians) to address a complex training challenge (sepsis recognition), speaks to the flexibility and power of these methods. Third, the Handbook of *Augmented Reality Training Design Principles* (Militello, Sushereba, Ramachandran, 2023) fills an important gap. Prior to this book, there were no resources consolidating relevant research and theory into design principles to guide developers of AR-based training. Lastly, ICA is a first-of-kind framework to guide the design of effective human-machine teams that will use advanced automation and autonomy in work contexts of the future (Ernst et al., 2019).

With regard to practical impact, I have been a leader in making CTA widely accessible via publications, workshops, and online self-study courses. I actively promulgated and exercised the DCD framework, increasing the accessibility of CTA methods for multidisciplinary teams.

Perhaps my most significant practical impact is contributing to sepsis recognition training that is in use today (Patterson et al., 2016), leading to improved outcomes for patients.

6.4 Future Work

Future work will continue to focus on adapting and applying CTA methods to address challenging real-world problems. As we continue to adapt and extend CTA methods, the CTA Institute provides a platform that will allow us to add training resources so they are available worldwide. In my role as Vice Executive Director of the NDM Association, I work to create a community that shares methodological innovations, builds on prior knowledge, and engages with practitioners and scientists outside of the NDM core. I aim to create a robust community so that NDM research continues beyond the work of its founders.

My research will continue to address societal problems. In the near term, I expect there to be considerable focus on designing to support humans and technology working together in the age of large language models, autonomy, and advanced automation. The ICA framework has been used successfully to explore envisioned worlds in Army Aviation and Mars missions. I will continue to apply the framework in other domains, exploring strategies to make it more pragmatically feasible to meet aggressive timelines and limited budgets.

I will also continue to focus on recognition skills training. After spending years writing a book on this topic, I have already discovered concepts I wish I had included. I intend to continue to explore how emerging technologies can be used to create training that keeps combat medics (and others that make decisions in the worst situations) safe and effective. Importantly, this will include strategies for evaluating training efficacy so that we can be confident the training improves cognitive skills as intended.

6.5 Summary and Conclusions

This synoptic report recounts 30+ years of applied research, shaped by societal needs. Scientific contributions are primarily methodological, exploring novel ways to study complex problems and invisible aspects of work. Practical contributions relate largely to making CTA methods accessible, and applying CTA to difficult problems such as sepsis recognition. Future work will focus on continuing to grow the NDM community as founders retire, extending CTA methods to meet emerging challenges associated with advanced autonomy and AI, and developing training for people who make decisions in high stakes environments.

REFERENCES

Angus, D. C., & Wax, R. S. (2001). Epidemiology of sepsis: An update. *Critical Care Medicine*, *29*(7), S109-S116. https://doi.org/10.1097/00003246-200107001-00035

Assadi, A., Laussen, P. C., Freire, G., Ghassemi, M., & Trbovich, P. (2022). Decision-centered design of a clinical decision support system for acute management of pediatric congenital heart disease. *Frontiers in Digital Health*, *4*, 1016522. https://doi.org/10.3389/fdgth.2022.1016522

Borders, J., & Klein, G. (2017, June). The critical decision audit: blending the critical decision method & the knowledge audit. In *Proceedings of the 13th bi-annual international conference on naturalistic decision making,* 33-9.

Borders, M., Militello, L., Sushereba, C. (2014). Colorectal Cancer Screening Decision Support: Phase 2 Final Report [Unpublished technical report, Contract No: 200-2012-52902]. Applied Decision Science.

Brown, O., Power, N., & Gore, J. (2024). Cognitive task analysis: Eliciting expert cognition in context. *Organizational Research Methods, 28*(3), 375-404. https://doi.org/10.1177/10944281241271216

Organizational learning and technological change (pp. 16-46). Springer. https://doi.org/10.1007/978-3-642-79550-3_2

Coffey, A., Beverley, H., & Paul, A. (1996). Qualitative data analysis: Technologies and representations. *Sociological research online*, *1*(1), 80-91. https://doi.org/10.5153/sro.1

Crandall, B. & Calderwood, R. (1989). *Clinical assessment skills of experienced neonatal intensive care nurses* [Unpublished technical report, Contract 1-R43-NR01911-01]. Klein Associates.

Crandall, B. & Gamblion, V. (1991). *Guide to early sepsis assessment in the NICU.* [Unpublished technical report, Contract 1-R43-NR01911-01]. Klein Associates.

Crandall, B. & Getchell-Reiter, K. (1993). Critical decision method: A technique for eliciting concrete assessment indicators from the “intuition” of NICU nurses. *Advances in Nursing Sciences, 16*(1), 42-51. https://doi.org/10.1097/00012272-199309000-00006

Crandall, B., Klein, G. A., & Hoffman, R. R. (2006). *Working minds: A practitioner's guide to cognitive task analysis*. MIT Press. https://doi.org/10.7551/mitpress/7304.001.0001

Dale, S. (2015). Heuristics and biases: The science of decision-making. *Business Information Review*, *32*(2), 93-99. DOI: 10.1177/0266382115592536

Dempsey, D., Sargent, R., Hobbs, A., Latorella, K., Zhaofeng, H., Niemann, C., Chaikin, A., Lewis, R., Robinson, P., Karasinksi, J., Tyson, T., Shear-Tungol, T., Durhan, M. Lippert, A., Landon, L., Hughes-Baird, S., Keyzer, K., Dempsey, M., Maddox, J., Pak, S., Null, C. (2024). Trade Space Analysis: Balancing Crew and Mission Design Parameters. (NASA/TM-20240011265 NESC-RP-20-01525). https://ntrs.nasa.gov/api/citations/20240011265/downloads/12-10-24%2020240011265.pdf

Department of the Army. (2019, July). *ADP 5-1 The Operations Process* (Army Doctrine Publication 5-1). https://rdl.train.army.mil/catalog-ws/view/100.ATSC/E4166A5D-0FE7-4780-916A-A7E9B227147C-1337689957702/adp5_0.pdf

Dyer, J.S. (2005). MAUT-Multiattribute Utility Theory. In J. Figueira, S. Greco, & M. Ehrogott (Eds.) *Multiple criteria decision analysis: State of the art surveys. International series in 0perations research & management science, 78*, (pp. 265-292). New York. https://doi.org/10.1007/0-387-23081-5_7.

Ernst, K., Militello, L., Roth, E., Sushereba, C., Diiulio, J., Scheff, S., Klein, D., & Wonderly, S. (in press). Integrated Cognitive Analysis for Human-Automation Function Allocation and System Design. *Journal of Aircraft.*

Ernst, K., Roth, E., Militello, L., Sushereba, C., DiIulio, J., Wonderly, S., Scheff, S., Klein, D., & Taylor, G. (2019). A strategy for determining optimal crewing in future vertical lift: Human-automation function allocation. *Proceedings of the Vertical Flight Society's 75th Annual Forum & Technology Display*, Vertical Flight Society, 1-15. https://doi.org/10.4050/f-0075-2019-14558

Ernst, K., Winner, J., Goldberg, B., Cheng, T., Plaza-Verdiun, M., Borders, J., & Militello, L. (2023, September). Training for High-Stakes Domains. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, *67*(1), 1356-1360. https://doi.org/10.1177/21695067231192413

Fitzgerald, G., Morris, R., Militello, L., & Fletcher, J. (2024, September). Building Trust in Human-Machine Teaming for Autonomous Space Sensing. *Proceedings of the 25th Advanced Maui Optical and Space Surveillance (AMOS) Technologies Conference,* 22-35. https://amostech.com/TechnicalPapers/2024/Machine-Learning-for-SDA/Fitzgerald.pdf

Geis, G.L., Wheeler, D.S., Bunger, A., Militello, L.G., Taylor, R.G., Bauer, J.P., Byczkowski, T.L., Kerrey, B.T., & Patterson, M.D. (2018). A validation argument for a simulation-based training course centered on assessment, recognition and early management of pediatric sepsis. *Simulation in Healthcare*, *13*(1), 16-26. *DOI:* 10.1097/SIH.0000000000000271

Glaser, R., & Resnick, L.B. (1972). Instructional psychology. *Annual Review of Psychology, 23*, 207-276. https://doi.org/10.1146/annurev.ps.23.020172.001231

Gualtieri, J., Elm, W., Easter, J., Hanes, L. F., & Naser, J. (2012) Decision-centered design guidelines for the electric power industry's human systems interfaces. *Proceedings of the 8th International Meeting on Nuclear Plant Instrumentation Control and Human-Interface Technologies*. American Nuclear Society.

Han, Y. Y., Carcillo, J. A., Dragotta, M. A., Bills, D. M., Watson, R. S., Westerman, M. E., & Orr, R. A. (2003). Early reversal of pediatric-neonatal septic shock by community physicians is associated with improved outcome. *Pediatrics*, *112*(4), 793-799. https://doi.org/10.1542/peds.112.4.793

Harle, C. A., DiIulio, J., Downs, S. M., Danielson, E. C., Anders, S., Cook, R. L., Hurley, R.W., Mamlin, B.W., & Militello, L. G. (2019). Decision-centered design of patient information visualizations to support chronic pain care. *Applied Clinical Informatics*, *10*(04), 719-728. https://doi.org/10.1055/s-0039-1696668

Hoffman, R.R., Crandall, B., & Shadbolt, N. (1998) Use of the Critical Decision Method to elicit expert knowledge: A case study in the methodology of cognitive task analysis. *Human Factors, 40*(2)*,* pp. 254-276. DOI: 10.1518/001872098779480442

Hoffman, R. & Militello, L.G. (2008). *Perspectives on Cognitive Task Analysis: Historical Origins and Modern Communities of Practice*. Taylor and Francis https://doi.org/10.4324/9780203809877

Hollnagel, E., & Woods, D. D. (1983). Cognitive systems engineering: New wine in new bottles. *International journal of man-machine studies*, *18*(6), 583-600. https://doi.org/10.1016/s0020-7373(83)80034-0

Huntington Ingalls Industries Mission Technologies. (2025). “Human Performance Modeling: IMPRINT,” retrieved 6 Aug. 2025.

Hutton, R. J., Miller, T. E., & Thordsen, M. L. (2003). Decision-centered design: Leveraging cognitive task analysis in design. In E. Hollnagel (Ed.), *Handbook of cognitive task design* (pp. 383-416). CRC Press. https://doi.org/10.1201/9781410607775

Johnston, D. & Morrison, B. W. (2016). The application of naturalistic decision-making techniques to explore cue use in rugby league playmakers. *Journal of Cognitive Engineering and Decision Making*, *10*(4), 391-410. https://doi.org/10.1177/1555343416662181.

Jonassen, D., Militello, L.G., & Crandall, B. (1998). Critical Incident/Critical Decision Method. In D.H. Jonassen, M. Tessmer, and WH. Hannum, *Task Analysis for Instructional Design*. (pp. 181-192). Lawrence Erlbaum. https://doi.org/10.4324/9781410602657-28

Kaempf, G.L., Wolf, S., Thordsen, M.L., & Klein, G. (1992). Decision making in the AEGIS combat information center. [Unpublished technical report, Contract N66001-90-6023]. Klein Associates.

Kahneman, D. (2011). *Thinking, fast and slow*. Macmillan. https://doi.org/10.1007/s00362-013-0533-y

Kaplan, A. D., Cruit, J., Endsley, M., Beers, S. M., Sawyer, B. D., & Hancock, P. A. (2020). The Effects of Virtual Reality, Augmented Reality, and Mixed Reality as Training Enhancement Methods: A Meta-Analysis. *Human Factors, 63*(4), 706-726. https://doi.org/10.1177/0018720820904229 (Original work published 2021)

Kirkpatrick, J. D., & Kirkpatrick, W. K. (2016). *Kirkpatrick's four levels of training evaluation*. Association for Talent Development.

Klein, G. (2009). *Streetlights and Shadows*. MIT Press. https://doi.org/10.1080/00140139.2014.962288

Klein, G., Calderwood, R, & MacGregor, D., (1989). Critical decision method for eliciting knowledge. *IEEE Transactions on Systems, Man, & Cybernetics, 19*(3), 462-472. https://doi.org/10.1109/21.31053

Klein, G., Hoffman, R. R., Clancey, W. J., Mueller, S. T., Jentsch, F., & Jalaeian, M. (2023). “Minimum Necessary Rigor” in empirically evaluating human–AI work systems. *AI Magazine*, *44*(3), 274-281. https://doi.org/10.1002/aaai.12108

Klein, G., & Militello, L. (2001). Some guidelines for conducting a cognitive task analysis. In E. Salas (Ed.), *Advances in human performance and cognitive engineering research* (pp. 163–199). Elsevier Science/JAI Press. https://doi.org/10.1016/S1479-3601(01)01006-2

Klein, G., Militello, L., Dominguez, C., & Lintern, G. (2017). A One-Day Workshop for Teaching Cognitive Systems Engineering Skills. In P.J. Smith & R.R. Hoffman (Eds.), *Cognitive Systems Engineering: The Future for a Changing World*. CRC Press. https://doi.org/10.1201/9781315572529-16

Klein, G.A., Orasanu, J., Calderwood, R., & Zsambok, C. E. (1993). *Decision making in action: Models and methods.* Ablex. https://drive.google.com/file/d/1Q4fdnGxmkLoZ_Tyg3RrMhxaLm8v6asvZ/view

Klein, G.A. & Thordsen, M.L. (1991). Representing cockpit crew decision making. In R.S. Jensen (Ed.) *Proceedings of the sixth International Symposium on Aviation Psychology*, 1026-1031. Ohio State University. https://books.google.com/books/about/Proceedings_of_the_Sixth_International_S.html?id=2SJ4twAACAAJ

Larcel, D. & Andrews, D. H. (2021). Cognitive task analysis of unmanned aircraft system pilots. *International Journal of Aerospace Psychology, 31*(4), 319-342. http://dx.doi.org.wrs.idm.oclc.org/10.1080/24721840.2021.1895797.

Lundberg, J. S., Perl, T. M., Wiblin, T., Costigan, M. D., Dawson, J., Nettleman, M. D., & Wenzel, R. P. (1998). Septic shock: An analysis of outcomes for patients with onset on hospital
wards versus intensive care units. *Critical Care Medicine*, *26*(6), 1020-1024. https://doi.org/10.1097/00003246-199806000-00019

Martini, N., Choong, J. W., Cruz, P. D. D., Lau, H., Lim, H., Liu, R., Lim, A. G., & Marshall, D. (2022). Assessing antibiotic prescribing in nurse practitioners: Applied cognitive task analysis. *International Journal of Nursing Studies Advances, 4*, 1-12. https://doi.org/10.1016/j.ijnsa.2022.100101

McDermott, P., Dominguez, C., Kasdaglis, N., Ryan, M., Trahan, I., & Nelson, A. (2018). Human-machine teaming systems engineering guide. *MITRE Corp Bedford MA Bedford United States.* https://www.mitre.org/sites/default/files/2021-11/prs-17-4208-human-machine-teaming-systems-engineering-guide.pdf

Militello, L.G., DiIulio, J.B., Borders, M.R., Sushereba, C.E., Saleem, J.J., Haverkamp D, Imperiale, T.F. (2017) Evaluating a modular decision support application for colorectal cancer screening. *Applied Clinical Informatics, 8*(1), 162-179. https://doi.org/10.4338/ACI-2016-09-RA-0152.

Militello, L.G. & Hutton, R.J.B. (1998). Applied Cognitive Task Analysis (ACTA): A practitioner's toolkit for understanding cognitive task demands. *Ergonomics Special Issue: Task Analysis*, *41* (11), 1618 -1641. https://doi.org/10.1080/001401398186108

Militello, L.G. & Klein, G. (2013). Decision-Centered Design. In J.D. Lee & A. Kirlik (Eds.). *The Oxford handbook of cognitive engineering* (pp. 261-271). Oxford University Press. https://doi.org/10.1093/oxfordhb/9780199757183.013.0016

Militello, L. G., Rattray, N. A., Flanagan, M. E., Franks, Z., Rehman, S., Gordon, H. S., Barach, P., & Frankel, R. M. (2018). "Workin' on our night moves": How residents prepare for shift handoffs. *The Joint Commission Journal on Quality and Patient Safety*, *44*(8), 485-493. https://doi.org/10.1016/j.jcjq.2018.02.005

Militello, L.G., Roth, E.M., Scheff. S., Ernst, K., Sushereba, C., Diiulio, J., Klein, D., Wonderly, S. (2019a). Crew configuration analysis for future airborne reconnaissance operations. [Unpublished technical report, Contract No. NNX16AJ91A, 21-1614-5637-ADS2018]. Applied Decision Science

Militello, L.G., Roth, E.M., Scheff. S., Ernst, K., Sushereba, C., Diiulio, J., Klein, D. (2019b). Toward an optimally crewed future vertical lift vehicle: Crewing strategies and recommendations. [Unpublished technical report, Contract No. NNX16AJ91A, 21-1614-5637-ADS2018]. Applied Decision Science.

Militello, L.G., Saleem. J.J., Borders, M.R., Sushereba, C.E., Haverkamp, D., Wolf, S.P., Doebbeling, B.N. (2016). Designing colorectal cancer screening decision support: A cognitive engineering enterprise. *Journal of Cognitive Engineering and Decision Making, 10*(1), 74-90. https://doi.org/10.1177/1555343416630875

Militello, L. G., Salwei, M. E., Reale, C., Sushereba, C., Slagle, J. M., Gaba, D., Weinger, M.B., Rask, J., Faiman, J., Andreae, M., Burden, A.R., Ander, S. (2023). Adapting cognitive task analysis methods for use in a large sample simulation study of high-risk healthcare events. *Journal of Cognitive Engineering and Decision Making*, *17*(4), 315-331. https://doi.org/10.1177/15553434231192283

Militello, L.G., Sushereba, C.E., Cheng, T., Kaduce, M., Read, J.M., Wagner, E., Smith, O.W., Goolsby, C., Winner, J. (2023 August 14-17.) *Error recovery training for hemorrhage control* [Poster presentation]. Military Health Services Research Symposium. Kissimmee, FL.Militello, L. G., Sushereba, C. E., & Ramachandran, S. (2023). *Handbook of augmented reality training design principles*. Cambridge University Press. https://doi.org/10.1017/9781009216166.001

Miller, M. J., & Feigh, K. M. (2019). Addressing the envisioned world problem: A case study in human spaceflight operations. *Design Science*, *5*(e3). https://doi.org/10.1017/dsj.2019.2

Miller, T. E., Stanard, T. W., & Wiggins, S. L. (2003, October). Applying decision-centered design to damage control for the Navy. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting (47)*3, 278-282). https://doi.org/10.1177/154193120304700306

Odetola, F. O., Rosenberg, A. L., Davis, M. M., Clark, S. J., Dechert, R. E., & Shanley, T. P. (2008). Do outcomes vary according to the source of admission to the pediatric intensive care unit?*. *Pediatric Critical Care Medicine, 9(1), 20-25.* https://doi.org/10.1097/01.pcc.0000298642.11872.29

Papautsky, E. L., Strouse, R., & Dominguez, C. (2020). Combining cognitive task analysis and participatory design methods to elicit and represent task flows. *Journal of cognitive engineering and decision making, 14*(4), 288-301. https://doi.org/10.1177/1555343420976014.

Patterson, M.D., Militello, L.G., Bunger, A., Taylor, R.G., Wheeler, D.S., Klein, G., Geis, G.L. (2016). Leveraging the critical decision method to develop simulation-based training for early recognition of sepsis, *Journal of Cognitive Engineering and Decision Making, 10*(1), pp. 36-56. https://doi.org/10.1177/1555343416629520

Patterson, E.S. & Miller, J.E. (2010). Macrocognition metrics and scenarios: Design and evaluation for real-world teams. CRC Press. https://doi.org/10.1201/9781315593173

Potter, S. S., Roth, E. M., Woods, D. D., & Elm, W. C. (2000). Bootstrapping multiple converging cognitive task analysis techniques for system design. In M. Schraagen, S.F. Chipman, and V.L. Shalin (Eds.) *Cognitive task analysis* (pp. 317-340). Lawence Erlbaum Associates.

Resnick, L.G. (1976). Task analysis in instructional design: Some cases from mathematics. In D. Klahr (Ed.), *Cognition and instruction* (pp. 51-80). Erlbaum. https://doi.org/10.21236/ada024717

Roth, E.M., Sushereba, C., Militello, L.G., Diiulio, J., Ernst, K. (2019). Function Allocation Considerations in the Era of Human Autonomy Teaming. *Journal of Cognitive Engineering and Decision Making* 13(4), pp. 199–220. https://doi.org/10.1177/1555343419878038

Saleem J.J., Haggstrom D.A., Militello L.G., Flanagan M., Kiess C.L., Arbuckle N., Doebbeling B.N. (2011). Redesign of a computerized clinical reminder for colorectal cancer screening: A human-computer interaction evaluation. *BMC Medical Information Decision Making, 11(*74), 1-11. https://doi.org/10.1186/1472-6947-11-74

Saleem, J. J., Militello, L. G., Arbuckle, N. B., Flanagan, M., Haagstrom, D., Linder, J.A., & Doebbeling, B. (2009). Clinical decision support at three benchmark institutions for health IT: Provider perceptions of colorectal cancer screening. *Proceedings of the American Medical Informatics Association Annual Symposium* (pp. 558–562). American Medical Informatics Association. https://www.proceedings.com/american-medical-informatics-association-amia/

Saleem, J. J., Patterson, E. S., Militello, L., Render, M. L., Orshansky, G., & Asch, S. A. (2005). Exploring barriers and facilitators to the use of computerized clinical reminders. *Journal of the American Medical Informatics Association*, *12*, 438–47. https://doi.org/10.1197/jamia.m1777

Shorrock, S. (2022). Adaptive Imagination at Work in Health Care. In H. Laroche, C. Bieder, J. Villena-Lopez (Eds.) *Managing Future Challenges for Safety* (pp. 95-104). Springer https://doi.org/10.1007/978-3-031-07805-7_12

Sidebottom, A. C., Collins, B., Winden, T. J., Knutson, A., & Britt, H. R. (2012). Reactions of nurses to the use of electronic health record alert features in an inpatient setting. *Computers Informatics Nursing*, *30*(4), 218–226. https://doi.org/10.1097/ncn.0b013e3182343e8f

Strauss, A., & Corbin, J. (1994). Grounded theory methodology: An overview. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 273–285). Sage.

Streiff, M. B., Carolan, H. T., Hobson, D. B., Kraus, P. S., Holzmueller, C. G., Demski, R., Lau, B.D., Biscup-Horn, P., Pronovost, P.J., Haut, E. R. (2012). Lessons from the Johns Hopkins multi-disciplinary venous thromboembolism (VTE) prevention collaborative. *BMJ, 344*, e3935. https://doi.org/10.1136/bmj.e3935

Sushereba, C., Fernandez, R., Plaza-Verduin, M., Grand, J., Wolf, S., Hendricks, H., & Militello, L. (2024). *Developing a Just-in-time Refresher Trainer for Advanced Life Support in Austere Regions: Phase II Final Report*. [Unpublished technical report, Contract No. W81XWH-22-C-0017]. Applied Decision Science, LLC.

Sushereba, C.E., Militello, L.G., Read, J., Wagner, E. (2024). *Error recovery training phase 2: Final Report*. [Unpublished technical report, Contract No. FA8650-20-C-6249]. Applied Decision Science, LLC.

Tawfik, H., Anya, O., & Nagar, A.K. (2012). Understanding clinical work practices for cross-boundary decision support in e-health. *IEEE Transactions on Information Technology in Biomedicine*, *16*(4), 530–554. https://doi.org/10.1109/titb.2012.2187673

Taynor, J., Crandall, B., & Wiggins, S. (1987). The reliability of the critical decision method [Unpublished technical report, Contract MDA903-86-C-0170]. Klein Associates.

Thordsen, M.L., Militello, L.G., & Klein, G.A. (1992). Cognitive task analysis of critical team decision making during multiship engagements [Unpublished technical report, Contract F33615-90-C-0005]. Klein Associates.

Wang, M. H., Militello, L. G., Morris, R., & Flosenzier, V. (2024, September). Decision-Centered Design for Post-implementation Analysis of Automation in the Petrochemical Industry. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 68(1), 95-97. https://doi.org/10.1177/10711813241260339

Woods, D., & Dekker, S. (2000). Anticipating the effects of technological change: A new era of dynamics for human factors. *Theoretical issues in ergonomics science*, *1*(3), 272-282. https://doi.org/10.1080/14639220110037452

Yates, K.A., & Feldon, D.F. (2011). Advancing the practice of cognitive task analysis: a call for taxonomic research. *Theoretical Issues in Ergonomics Science, 12*(6), 472-495. https://doi.org/10.1080/1463922x.2010.505269

LIST OF APPENDICES

APPENDIX 1.1 TIMELINE OF KEY PUBLICATIONS

Shading indicates the most relevant theme Blue = CTA development (Chapter 2); Orange = training applications (Chapter 3); Green = technology design (Chapter 4)

Year	Publication
ongoing	Error Recovery Training for Hemorrhage Control (Sushereba et al., in preparation)
2024	Get on the Round Dial: Fighter pilot strategies for recovering situational awareness after disorienting physiological events (Militello et al., 2024)
2023	Adapting CTA methods for use in large sample simulation study of high-risk healthcare events (Militello, Salwei et al., 2023)
	Handbook of augmented reality training design principles. (Militello, Sushereba et al., 2023)
2022	
2021	Error recovery training literature review: implications for emergency field medicine. (Militello et al., 2021)
2020	Strategems: Embedding cognitive training in game-based environments (Newsome et al., 2020)
	Virtual patient immersive trainer to train perceptual skills using augmented reality (Sushereba & Militello, 2020)
	Primary care clinician’s beliefs and strategies for managing chronic pain in an era of a national opioid epidemic (Militello et al., 2020)
2019	A strategy for determining optimal crewing in future vertical lift: Human-automation function allocation (Militello et al., 2019)
	Adapting cognitive task analysis to investigate clinical decision-making and medication safety incidents (Russ et al., 2019)

Year	**Publication**
	Crew configuration analysis for future airborne reconnaissance operations (Militello et al., 2019)
	Envisioning user requirements for first-of-a-kind future rotorcraft (Sushereba et al., 2019)
2018	“Workin’ on our night moves”: how residents prepare for shift handoffs. (Militello, Rattray et al., 2018)
	Understanding how primary care clinicians make sense of chronic pain (Militello, Anders et al., 2018)
	Toward an optimally crewed future vertical lift vehicle: Overview of the envisioned world, core missions, and pertinent technology (Militello, Roth et al., 2018)
	A validation argument for a simulation-based training course centered on assessment, recognition and early management of pediatric sepsis (Geis et al., 2018)
2017	Evaluating a modular decision support application for colorectal cancer screening (Militello et al., 2017)
	A One-Day Workshop for Teaching Cognitive Systems Engineering Skills (Klein et al., 2017)
2016	Designing colorectal cancer screening decision support: A cognitive engineering enterprise (Militello et al., 2016)
	Leveraging the critical decision method to develop simulation-based training for early recognition of sepsis (Patterson et al., 2016)
2015	Designing for military pararescue: Naturalistic decision-making perspective, methods, and frameworks. (Militello et al., 2015)
2014	
2013	Decision-Centered Design (Militello & Klein, 2013)
2012 - 2009	

Year	Publication
2008	Perspectives on cognitive task analysis: Historical origins and modern communities of practice (Hoffman & Militello, 2008)
1999-2007	
1998	Applied Cognitive Task Analysis (ACTA) A practitioner’s toolkit for understanding cognitive task demands (Militello & Hutton, 1998)
	Critical Incident/Critical Decision Method (Jonassen et al., 1998)

Full References for articles not cited in the body of the thesis

Militello, L. G., Anders, S., Downs, S. M., Diiulio, J., Danielson, E. C., Hurley, R. W., & Harle, C. A. (2018). Understanding how primary care clinicians make sense of chronic pain. *Cognition, Technology & Work*, *20*(4), 575-584. https://doi.org/10.1007/s10111-018-0491-1

Militello, L. G., Ernst, K. M., Panganiban, A. R., Wagner, E., Tolston, M., & Funke, G. J. (2024). Get on the Round Dial: Fighter Pilot Strategies for Recovering Situation Awareness After Disorienting Physiological Events. *Journal of Cognitive Engineering and Decision Making, 18*(2), 137-155. https://doi.org/10.1177/15553434241234105

Militello, L.G., Hurley, R. W., Cook, R. L., Danielson, E. C., Diiulio, J., Downs, S. M., … Harle, C. A. (2020). Primary care clinicians’ beliefs and strategies for managing chronic pain in an era of a national opioid epidemic. *Journal of General Internal Medicine, 35*, 3542–3548. https://doi.org/10.1007/s11606-020-06178-2

Militello, L.G., Roth, E.M., Scheff. S., Ernst, K., Sushereba, C., Diiulio, J., Klein, D., Wonderly, S. (2019). Crew configuration analysis for future airborne reconnaissance operations. [Unpublished technical report, Contract No. NNX16AJ91A, 21-1614-5637-ADS2018, Applied Decision Science.

Militello, L.G., Sushereba, C.E., Branlat, M., Bean, R., Finomore, V. (2015). Designing for military pararescue: Naturalistic decision-making perspective, methods, and frameworks. *Journal of Occupational and Organizational Psychology*. *88*, 251-272. doi:10.1111/joop.12114. https://doi.org/10.1111/joop.12114

Militello, L. G., Wagner, E., Winner, J., Sushereba, C., & McCool, J. (2021). *Error recovery training literature review: Implications for emergency field medicine*. *Proceedings of the 65th Human Factors and Ergonomics Society Annual Meeting, 65*, 495–499. Los Angeles: Sage. https://doi.org/10.1177/1071181321651073

Newsome, E., Militello, L., & Ramachandran, S. (2020). Stratagems: Embedding Cognitive Training in Game-Based Environments. In *Advances in Usability, User Experience, Wearable and Assistive Technology: Proceedings of the AHFE 2020 Virtual Conferences on Usability and User Experience, Human Factors and Assistive Technology, Human Factors and Wearable Technologies, and Virtual Environments and Game Design, July 16-20, 2020, USA* (pp. 589-595). Springer International Publishing. https://doi.org/10.1007/978-3-030-51828-8_77

Sushereba, C. E., & Militello, L. G. (2020). Virtual patient immersive trainer to train perceptual skills using augmented reality. *Proceedings of the 64th Human Factors and Ergonomics Society Annual Meeting, 64*, 470–472. Los Angeles: Sage. https://doi.org/10.1177/1071181320641106

Sushereba, C., Militello, L., Ernst, K., DiIulio, J., Roth, E., Scheff, S., Huff, W., IV. (2019). Envisioning user requirements for first-of-a-kind future rotorcraft. Paper presented at the 20th International Symposium on Aviation Psychology, Dayton, OH. https://doi.org/10.4050/f-0075-2019-14558

APPENDIX 3.1 ACTA APPLICATIONS

	Participants	Application
Healthcare		
Green, L. A., Potworowski, G., Day, A., May-Gentile, R., Vibbert, D., Maki, B., & Kiesel, L. (2015). Sustaining "meaningful use" of health information technology in low-resource practices. *Annals of Family Medicine*, 13, 17-22. https://doi.org/10.1370/afm.1740.	Clinicians and EMR implementation specialists	Identify barriers to EMR maintenance in low-resource practices
Hildebrand, E. A., Branaghan, R. J., Hallbeck, M. S., & Blocker, R. C. (2014). Team briefings in the gynecological operating room: A cognitive task analysis. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 58, 753-757. https://doi.org/10.1177/1541931214581137.	Surgical teams	Team briefing model
Ma, J. & Drury, C. G. (2004). Cognitive task analysis of data mining processes in bioinformatics research. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 48(5).	Bioinformatics researchers	Data mining techniques
Martini, N., Choong, J. W., Cruz, P. D. D., Lau, H., Lim, H., Liu, R., Lim, A. G., & Marshall, D. (2022). Assessing antibiotic prescribing in nurse practitioners: Applied cognitive task analysis. *International Journal of Nursing Studies Advances*, 4, 1-12. https://doi-org.wrs.idm.oclc.org/10.1016/j.ijnsa.2022.100101.	Nurse practitioners	Training
Marzullo, J., Farahani, A., Fendley, M., & Caldorera-Moore, M. (2022). Incorporation of the ACTA method for bone defect interpretation training. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 66, 412-416. https://doi.org/10.1177/1071181322661348.	Radiologists	Training/Technology design
Militello, L. G., Saleem, J. J., Borders, M. R., Sushereba, C. E., Haverkamp, D., Wolf, S. P., & Doebbeling, B. N. (2016). Designing colorectal cancer screening decision support: A cognitive engineering enterprise.	Primary-care physicians	Technology design

	Participants	Application
Journal of Cognitive Engineering and Decision Making, 10, 74-90. https://doi.org/10.1177/1555343416630875.		
Pfaff, M. S. (2015). Identifying option awareness requirements for public health crisis decision making. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 59, 548-552. https://doi.org/10.1177/1541931215591119.	Public health personnel	Training/Technology design/Interface design
Pickup, L., Lang, A., Shipley, L., Henry, C., Carpenter, J., McCartney, D., Butler, M., Hayes-Gill, B., & Sharkey, D. (2019). Development of a clinical interface for a novel newborn resuscitation device: Human factors approach to understanding cognitive user requirements. *JMIR Human Factors*, 6(2), 1-13. https://doi.org/10.2196/12055.	Neonatal Clinicians	Interface design
Shackak, A., Hadas-Dayagi, M., Zix, A., & Reis, S. (2009). Primary care physicians' use of an electronic medical record system: A cognitive task analysis. *Journal of General Internal Medicine*, 24(3), 341-348. doi: 10.1007/s11606-008-0892-6.	Physicians	Technology design
Military		
Capiola, A., Baxter, H. C., Pfahler, M. D., Calhoun, C. S., & Bobko, P. (2020). Swift trust in ad hoc teams: A cognitive task analysis of intelligence operators in multi-domain command and control. *Journal of Cognitive Engineering and Decision Making*, 14(3), 218-241. https://doi.org/10.1177/1555343420943460.	Intelligence operators	Antecedents to swift trust
Hutchins, S. G., Pirolli, P., & Card, S. (2003). Use of critical analysis method to conduct a cognitive task analysis of intelligence analysts. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 47(3), https://doi.org/10.1177/154193120304700348.	Intelligence analysts	Training

	Participants	Application
Larcel, D. & Andrews, D. H. (2021). Cognitive task analysis of unmanned aircraft system pilots. *International Journal of Aerospace Psychology*, 31(4), 319-342. http://dx.doi.org.wrs.idm.oclc.org/10.1080/24721840.2021.1895797.	Unmanned aircraft system pilots	Training
Minotra, D. & Feigh, K. (2017). Eliciting knowledge from helicopter pilots: Recommendations for revising the ACTA method for helicopter landing tasks. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 61, 242-246. https://doi.org/10.1177/1541931213601544.	Pilots	Technology design
Papautsky, E. L., Strouse, R., & Dominguez, C. (2020). Combining cognitive task analysis and participatory design methods to elicit and represent task flows. *Journal of cognitive engineering and decision making*, 14(4), 288-301. https://doi.org/10.1177/1555343420976014.	Submarine watch standers	Training
Tusl, M., Rainieri, G., Fraboni, F., DeAngelis, M., Depolo, M., Pietrantoni, L., & Pingitore, A. (2020). Helicopter pilots' tasks, subjective workload, and the role of external visual cues during shipboard landing. *Journal of Cognitive Engineering and Decision Making*, 14(3), 242-257. https://doi.org/10.1177/1555343420948720	Helicopter pilots	Technology design/Interface design
Information Technology		
Alarcon, G. M., Militello, L. G., Ryan, P., Jessup, S. A., Calhoun, C. S., & Lyons, J. B. (2017). A descriptive model of computer code trustworthiness. *Journal of Cognitive Engineering and Decision Making*, 11(2), 107-121. https://doi.org/10.1177/1555343416657236.	Programmers	Descriptive model
Gorman, M. G., Militello, L. G., Swierenga, S. J., & Walker, J. L. (2004). Internet searching by ear: Decision flow diagrams for sightless internet users. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 48(2), 243-247. https://doi.org/10.1177/154193120404800204.	Blind internet users	Interface design

	Participants	**Application**
Kleij, R. V., Schraagen, J. M., Cadet, B., & Young, H. (2022). Developing decision support for cybersecurity threat and incident managers. *Computer & Security*, 113, 1-15. https://doi.org/10.1016/j.cose.2021.102535.	IT security; Programming; Digital forensics; intelligence	Technology design/Interface design
Business & Finance		
McAndrew, C. & Gore, J. (2013). Understanding preferences in experience-based choice. *Journal of Cognitive Engineering and Decision Making*, 7(2), 179-197. https://doi.org/10.1177/1555343412463922.	Financial day traders	Decision theory/models
Osland, J., & Oddou, G., Bird, A., Osland, A. (2013). Exceptional global leadership as cognitive expertise in the domain of global change. *European Journal of International Management*, 7(5), 517-534. https://doi.org/10.1504/EJIM.2013.056475.	CEOs; COOs; EVP; VPs; Director of HR; Director of sales	Decision theory/models
Commercial Aviation		
Volz, K. M. & Dorneich, M. C. (2020). Evaluation of cognitive skill degradation in flight planning. *Journal of Cognitive Engineering and Decision Making*, 14(4), 263-287. https://doi.org/10.1177/1555343420962897.	Pilots; Flight instructors	Safety
Firefighting		
Nauert, E. & Gillan, D. (2019). The cognitive process of wildland fire chainsaw troubleshooting: Structure, content, and training. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 63, 1676-1680. https://doi.org/10.1177/1071181319631536.	Firefighters	Training/Technology design/Interface design
Hospitality		

	Participants	Application
Gore, J. & Riley, M. (2000). A study of the perceptions of the labour market by human resource managers in the UK hotel industry: A cognitive approach. *Tourism and Hospitality Research*, 2(3), 232-241. https://doi.org/10.1177/146735840000200304.	Human resource managers	Personnel recruitment
Offshore oil		
Hogenboom, S., Vinnem, J. E., Utne, I. B., & Kongsvik, T. (2021). Risk-based decision-making support model for offshore dynamic positioning operations. *Safety Science*, 140, 1-14. https://doi.org/10.1016/j.ssci.2021.105280.	Dynamic positioning operators	Training/Technology design
Manufacturing		
Seidelman, W., Lee, M., Kent, T. M., Carswell, C. M., Fu, B., & Yang, R. (2014). Development of a hybrid reality display for welders through applied cognitive task analysis. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 58, 1174-1178. http://dx.doi.org/10.1177/1541931214581245.	Welders	Technology design/Interface design
Sport		
Johnston, D. & Morrison, B. W. (2016). The application of naturalistic decision-making techniques to explore cue use in rugby league playmakers. Journal of Cognitive Engineering and Decision Making, 10(4), 391-410. https://doi.org/10.1177/1555343416662181.	Rugby league players	Training
Education		
Yusoff, N. M. & Salim, S. S. (2012). Investigating cognitive task difficulties and expert skills in e-learning storyboards using a cognitive task analysis technique. *Computers and Education*, 58, 652-665. https://doi.org/10.1016/j.compedu.2011.09.009.	Designers	Training

APPENDIX 4.1 BOOK REVIEWS

Appeared in: ***Human Factors and Ergonomics Society Bulletin*****, April 2024**

Book Review: Handbook of Augmented Reality Training Design Principals

By Laura G. Militello, Christen E. Sushereba, Sowmya Ramachandran

Upon first seeing the book one may think that it is way too thin to be worth the money, but you would be wrong. In two words, this book is complete and concise. The authors, all experienced in a variety of disciplines including cognitive and physical systems engineering, computer science, and training and learning, have put together a master text that any novice or expert in the field of training, learning, or technology can use today to support the effective employment of augmented reality technology with confidence. Having spent almost 30 years as a student, user, and applied practitioner in the development of expertise using training aids, devices, simulators, and simulations, I was apprehensive that such a short text could do justice to the space identified in the title. Edited volumes on the development of expertise generally are hundreds of pages thick. ***I was wrong***. The authors have concisely captured the essential elements from the critical areas of expertise development, cognitive systems engineering, metacognition, AR technology, and skill development through training and learning to produce a handbook that I have promoted to both new students, those proficient in the field of training, and those working in industry selling the technologies discussed.

One of the characteristics of the book that impressed me was the authors use of a wide range of foundational literature sources. Many new books covering established domains tend to neglect foundational literature which these authors do not. While their experiences have strongly influenced the writing, they have not made the book about their research, rather they have created a true handbook that anyone can use as a refresher or foundational text to teach and learn from. Eleven principals are laid out in the book. Each is explained in detail. Chapters begin with a discussion of the conditions that a principal applies to with pointers to the appropriate literature. Principals are then introduced followed by sections that discuss the principals and the what, why, when, and how of the AR technology and desired outcomes. Effectively weaved prose, figures, and structure create chapters that employ some of the most effective learning techniques that the authors discuss in the work itself. A consistent pattern allows the book to be useful in parts as a tool for focused training and learning or as a comprehensive discussion of the topic space.

Finally, the authors use real world examples to bring the theory to life, which highlights the absolute value of this work as one that should be on the bookshelf of anyone who considers themselves a practitioner or expert in the space of training and learning using augmented, virtual, or any other form of simulation technology.

Appeared in the *Society for Simulation in Europe (SESAM) Newsletter*, April 2024

Book review by Marc Lazarovici, SESAM Member and Past President

Review of "Handbook of Augmented Reality Training Design Principles" by Laura Militello, Christen E. Sushereba and Sowmya Ramachandran

Let me start by saying that I very much enjoyed reading this book – although it is, in fact, a textbook, it can be read almost like a good novel. The flow of ideas is very clear and in a certain way addictive, so you feel compelled to continue reading. While this might sound like a threat to some, it really is high praise.

The book introduces the reader to eleven design principles for creating training based on or using AR. The real-world examples are taken from a variety of fields, mainly healthcare and military, but all have in common the necessity for quick situation assessment and decision making. Thus, even for readers with a healthcare background, all examples are very easy to follow.

What sets this book apart is its versatility. Basically, you can go through the book chapter by chapter, in the order proposed by the authors – and that will create a coherent image on the current state of training using Augmented Reality (AR), along with the underlying design principles. However, the book's utility extends beyond a first-time read; it serves as a valuable reference for future consultations, allowing readers to delve directly into specific chapters or sections as needed. All presented concepts are well illustrated and supported by literature findings – so anyone interested in diving deeper into a specific field can easily do so by following the bibliography.

Moreover, the book does an excellent job of covering key components such as engagement, fidelity, and realism. It provides readers with a solid theoretical foundation while also exploring the practical aspects of scenario-based training in augmented reality. The concepts of scenario-based training in the context of augmented reality are well presented and help the reader, even if familiar with traditional simulation, to expand their horizon.

A particularly noteworthy feature of the book is the "Learn, Experience, Reflect" framework. This hands-on guide to applying the principles of AR in real-life training scenarios offers a clear, structured approach, and it ensures that the theoretical knowledge gained from the book can be effectively translated into practical application.

In conclusion, this textbook is a remarkable and useful resource for anyone interested in the intersection of education, technology, and augmented reality. Whether one is seeking to enhance their understanding of AR in training or looking for a reliable reference for future projects, I would definitely recommend it.

Appeared in the *Journal of Cognitive Engineering and Decision Making*, 2024, Vol 0-0

Book Review

Journal of Cognitive Engineering
and Decision Making
2024, Vol. 18(4) 426–428

DOI: 10.1177/15553434241283122
journals.sagepub.com/home/edm
Sage

Laura G. Militello, Christen E. Sushereba and S. Ramachandran (Eds). (2023) *Handbook of Augmented Reality Training Design Principles*. Cambridge, UK: Cambridge University Press

Reviewed by: Jan Maarten Schraagen,
Department Human Machine Teaming, Soesterberg, The Netherlands

The word 'Handbook' may conjure up images of hefty volumes covering entire academic fields, with dozens of chapters written by a wide range of experts. The aim of these types of handbooks is to provide an up-to-date and broad coverage of a particular area of interest and as such they have their place in the academic world. However, there is another meaning to the word 'handbook', one largely forgotten or ignored it seems, that emphasizes practical guidance, conciseness, and ease of accessibility. In this sense of the word, a handbook should be a 'handy book', one that distills practical advice from a large body of scientific research as well as practical experience by the authors. The Handbook of Augmented Reality Training Design Principles is such a handy book. It does not purport to offer a comprehensive overview of various theoretical perspectives on training design, nor does it offer up-to-date overviews of various augmented reality technologies. Rather, this book is written from a single theoretical perspective, that of Naturalistic Decision Making (Klein et al., 2010), or, phrased more broadly, macrocognition (Schraagen et al., 2008). The authors are world-leading experts in the fields of Naturalistic Decision Making and simulation-based training and their extensive experience with applying these fields to dynamic, high-stakes domains such as health care and military operations is apparent on every page.

Based on scientific literature and lessons learned from existing training, the authors offer eleven design principles to guide the design of augmented reality–based recognition skills training. The principles are grouped into five categories: engagement, scenario building, fidelity and realism, mental model construction, and scaffolding and reflection. These categories provide the organizing structure for this handbook and constitute its core body of chapters. These five chapters are preceded by an Introduction and a general chapter on recognition skills, and they are followed by a Synthesis chapter, explaining the Learn, Experience, Reflect framework to guide application of the design principles. The concluding chapter provides a concise summary of the design principles and a discussion of boundary conditions and challenges for augmented reality. Furthermore, each of the five core chapters discussing the principles is structured in an identical fashion, adding to ease of accessibility. First, the principle is explained, and its importance is argued for. Next, examples and empirical support are provided to illustrate each principle and provide their scientific evidence base. Subsequently, links to a macrocognition framework are provided, particularly with respect to three components of recognition skills: knowing what to attend to, creating meaning, and evaluating. This is followed by a brief summary and discussion and practical implications for training design. The handbook as a whole is concise at 136 pages, excluding references and index, and offers a lot of practical guidance.

(Report No. ADA199492). Klein Associates. https://apps.dtic.mil/sti/pdfs/ADA199492.pdf

Klein, G., Calderwood, R., & Clinton-Cirocco, A. (2010). Rapid decision making on the fire ground: The original study plus a postscript. *Journal of Cognitive Engineering and Decision Making*, *4*(3), 186–209. https://doi.org/10.1518/155534310x12844000801203

Militello, L. G., Sushereba, C. E., & Ramachandran, S. (2023). *Handbook of augmented reality training design principles*. Cambridge University Press.

Schraagen, J. M. C., Militello, L. G., Ormerod, T., & Lipshitz, R. (Eds.), (2008). *Naturalistic decision making and macrocognition*. Ashgate.

APPENDIX 6.1 REFLECTIVE ACCOUNT

In reflecting on my research journey, I am struck by the challenge of articulating research questions in the context of applied work. For many of the projects described in this synoptic report, sponsors came with real-world problems such as

- *Resident physicians are completing their training and unable to recognize sepsis,*
- *Navy instructors vary widely in how well they address the cognitive aspects of work in training,*
- *Designs for future rotorcraft that will fly faster and further than anything that exists today are not considering the principles of naturalistic decision making. (Will pilots be able to manage the cognitive workload in this envisioned world?)*

An important part of my role has been to work with sponsors to refine these needs into research questions and to frame and scope projects in meaningful ways. For me, this always involves searching the literature for methods and approaches that can be adapted and reaching out to researchers who have conducted prior work in the area. Furthermore, for these important applied problems, a well-functioning research team with complementary experiences and expertise is critical to success. For me, one of the hardest parts of writing this synoptic report has been trying to tease out what exactly was my contribution separate from the team. Although I did my best to estimate percent contributions in Table 1.1, I still find this exercise troubling.

The act of writing a synoptic summary has been a wonderful opportunity to take stock. When I look back, the early projects met with a fair amount of pushback. Reviewers of journal articles were unfamiliar with CTA methods and qualitative methods in general. I often received criticisms about sample size, inter-rated reliability, and different aspects of validity. At the time these critiques seemed unfair and irrelevant (due to the hubris of youth, no doubt), but they have in fact helped shape my thinking about what constitutes rigor for applied research using CTA and related methods. The enthusiasm of the practitioners we intended to support (emergency department physicians, Navy instructors, helicopter pilots, etc.) kept me going.

Conducting principled applied research requires one to straddle the operational and research communities. I will comment on a few challenges. One challenge is determining how many participants to include in a CTA study, what constitutes an expert, and whether to include participants at different skill levels. People frequently contact me asking advice on these issues, often looking for hard and fast rules. In truth, these decisions must be made based on the goals of the study and the pragmatic constraints of time, resources, and access to people who have the skills and experience needed. For example, if access to experts is limited, or does not exist yet in the case of envisioned worlds, we explore analogs: What other jobs have similar challenges and work context? We may conduct a few interviews and determine these are not the right people, sending us back to the drawing board to determine who has the types of expertise we need. We may conduct 5 interviews and determine we are not learning anything new, perhaps inspiring us to revisit the interview guide (are we asking the right questions?) or cut short the number of interviews (perhaps this aspect of work is not as complex as we initially thought).

A second challenge is managing the oversight process for meeting ethical guidelines for conducting human subjects research while conducting meaningful research. Because we work with universities, hospital systems, and government agencies, we are continuously adapting practices to meet the oversight procedures of our partner institutions. The biggest challenge is typically the requirement to specify exactly how the research will be conducted before we have begun. For exploratory, qualitative research this creates additional barriers to improving the methods and approach as the team learns more. I do my best to negotiate project extensions to accommodate the need for Institutional Review Board (IRB) amendments. To be clear, the challenge is not in maintaining ethical research practices. CTA studies rarely, if ever, raise even minor concerns for the principles of respect for persons, beneficence, and justice outlined in the Belmont report, all of which are carefully considered for each project. However, managing the lengthy multidisciplinary IRB processes is a significant challenge to project schedules and costs. I always adhere to the IRB requirements, but sometimes the research quality is constrained by our inability to adapt the methods without lengthy protocol amendment reviews.

A third challenge is disseminating findings. Although high quality journals such as the *Journal of Cognitive Engineering and Decision Making*, and *Cognition, Technology, & Work* are outlets for applied research using qualitative methods, these journals did not exist in the early 1990s. Many longer established journals are now open to publishing qualitative research, but this was not always the case. Earlier in my career, publication options were limited primarily to edited volumes and conference proceedings. I am happy to have seen and been a part of this shift. I served as an associate editor for the *Journal of Cognitive*

Engineering and Decision Making for eight years and remain a member of the editorial board.

A fourth challenge is transitioning research products to use. I am particularly proud of the sepsis recognition training we developed for Cincinnati Children's hospital and the AR-based training products Unveil provides to civilian and military medics because they are actually *in use*. Many of the prototype technologies, user interfaces, and training products I have developed and evaluated never made it to use because of bureaucratic barriers, limited funding, and my lack of knowledge and access to the right people. The gulf between research and applications is well documented and as difficult to overcome as ever.

Looking forward, I plan to continue to adapt CTA methods, particularly to meet the challenges of designing for safe and effective advanced automation and artificial intelligence. I will continue my quest to make CTA methods available to other applied researchers and to bridge the gap between the research and operational communities.