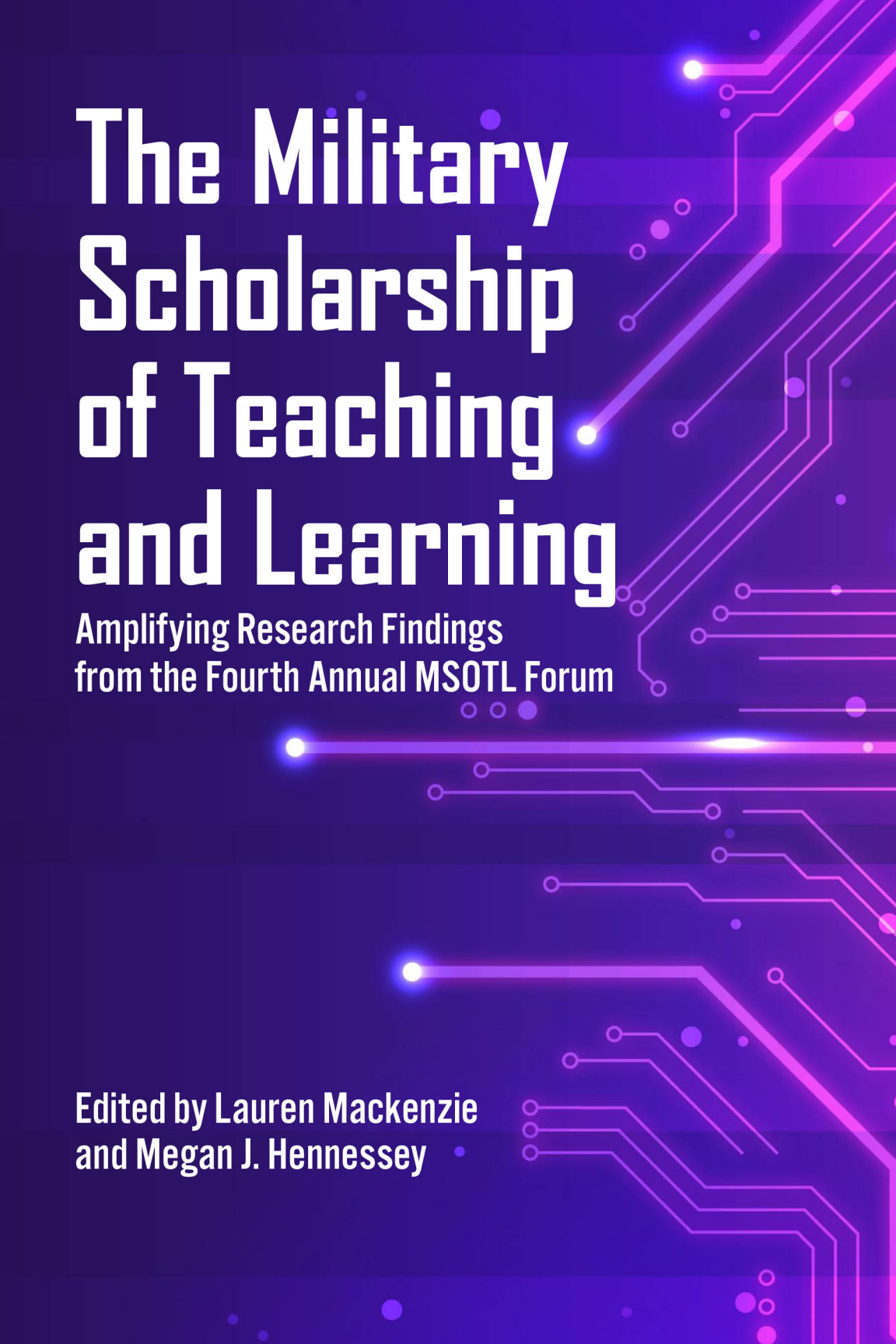




The Military Scholarship of Teaching and Learning

Amplifying Research Findings
from the Fourth Annual MSOTL Forum

Edited by Lauren Mackenzie
and Megan J. Hennessey

THE MILITARY SCHOLARSHIP OF TEACHING AND LEARNING



The Military Scholarship of Teaching and Learning

**Amplifying Research Findings
from the Fourth Annual MSOTL Forum**

edited by Lauren Mackenzie, PhD, and Megan J. Hennessey, PhD

MCUP

MARINE CORPS UNIVERSITY PRESS

Quantico, Virginia

2025

LIBRARY OF CONGRESS CATALOGING-IN-PUBLICATION DATA

Names: MSOTL Forum (4th : 2023 : Marine Corps University; Online), author. | Hennessey, Megan J., editor. |

Mackenzie, Lauren, 1976– editor. | Marine Corps University (U.S.). Press, issuing body.

Title: The military scholarship of teaching and learning : amplifying research findings from the fourth annual MSOTL Forum / edited by Megan Hennessey and Lauren Mackenzie.

Other titles: Amplifying research findings from the fourth annual MSOTL Forum

Description: Quantico, Virginia : Marine Corps University Press, 2025. | In scope of the U.S. Government Publishing Office Cataloging and Indexing Program (C&I); Federal Depository Library Program (FDLP) distribution status to be determined upon publication. | Includes bibliographical references and index. | Summary: “The Military Scholarship of Teaching and Learning (SOTL) field offers meaningful methods for educators to advocate for evidence-informed teaching, learning, and curriculum development on a global scale. The chapters included in this edited volume highlight such practices and measures and, in doing so, create the kind of educational synergy that has been the hallmark of the MSOTL community since its first informal gathering in 2019. This initiative expanded the traditional SOTL emphasis on/by civilian universities to create space for a military perspective in this growing field. This edited volume represents a further broadening of the relevance and applicability of the scholarship of teaching and learning for the military educator. The chapters illustrate a unique set of academic disciplines (ranging from meteorology to military history), methodologies (qualitative and quantitative), and learning levels (ranging from the community colleges to top-level schools). These 13 chapters highlight the continued innovation undertaken by professionals from across the continuum of military learning to excel both in the content they teach and the ways in which they teach it”— Provided by publisher.

Identifiers: LCCN 2024055919 (print) | LCCN 2024055920 (ebook) | ISBN 9798987336243 (paperback) | ISBN 9798987336250 (epub)

Subjects: LCSH: Military education—United States—Congresses.

Classification: LCC U405 .M76 2025 (print) | LCC U405 (ebook) | DDC 355.007--dc23 | SUDOC D 214.513:M 59

LC record available at <https://lccn.loc.gov/2024055919>

LC ebook record available at <https://lccn.loc.gov/2024055920>

Disclaimer

The views expressed in this publication are solely those of the authors. They do not necessarily reflect the opinion of Marine Corps University, the U.S. Marine Corps, the U.S. Navy, the U.S. Army, the U.S. Air Force, the U.S. Coast Guard, the U.S. Space Force, or the U.S. government. The information contained in this book was accurate at the time of printing. Every effort has been made to secure copyright permission on excerpts and artworks reproduced in this volume. Please contact the editors to rectify inadvertent errors or omissions. In the chapters that contain data obtained from human subject research, the authors have obtained permission from their institutions' internal research board and in alignment with the 1964 Declaration of Helsinki.

Copyright for all works—journals and monographs—is retained by the author(s). Some works of authorship from Marine Corps University Press (MCUP) are created by U.S. government employees as part of their official duties and are now eligible for copyright protection in the United States; further, some of MCUP's works available on a publicly accessible website may be subject to copyright or other intellectual property rights owned by non-Department of Defense (DOD) parties. Regardless of whether the works are marked with a copyright notice or other indication of non-DOD ownership or interests, any use of MCUP works may subject the user to legal liability, including liability to such non-DOD owners of intellectual property or other protectable legal interests.

MCUP products are published under a Creative Commons NonCommercial-NoDerivatives 4.0 International (CC BY - NC-ND 4.0) license.

Published by

Marine Corps University Press
2044 Broadway Street
Quantico, VA 22134

1st Printing, 2025

ISBN: 979-8-9873362-4-3

DOI: 10.56686/9798987336243

THIS VOLUME IS FREELY AVAILABLE AT WWW.USMCU.EDU/MCUPRESS



CONTENTS

Foreword	ix
Preface	xiii
Glossary of Terms	xv

Introduction	3
<i>By Lauren Mackenzie, PhD</i>	

Part I. Community & Creativity

Chapter 1	9
From Mindset and Collaboration to Connections and Experience: Creating a Teaching and Learning Community of Practice in the U.S. Air Force Leadership and Innovation Institute <i>By John Hinck, PhD; Steven Davis, PhD; Robert Hinck, PhD; Sara Kitsch, PhD; Mary Bartlett, PhD; and Andy Clayton, EdD</i>	

Chapter 2	47
What Do Creative Problem-Solvers Need? <i>By Lieutenant Colonel Timothy A. Sparks (USMC), EdD</i>	

Chapter 3	84
Creative Community of Learners—the Auxiliary Engine to Andragogical Approaches: Case Studies from the French War College <i>By Emilie Cleret and Jerome Collin</i>	

Part 2. Training & Education Case Studies

Chapter 4	101
What Does English Have to Do with War?: Decreasing Writing Error Propensity in a Military Service Academy First-Year Composition Course <i>By Lieutenant Commander Philip Garrow (USN), EdD</i>	

Chapter 5	120
Closing the Gap between Intellect and Character in an Online Ethics Course	
<i>By Colonel Heather Smigowski (USA), PhD</i>	
Chapter 6	136
Localized Professional Development: Creating Base-Wide Solutions to a Military-Wide Problem	
<i>By Layton Graves, EdD</i>	
Chapter 7	157
Making the Case against Unimodal Instructional Strategies: A Qualitative Intrinsic Case Study within Instructor Upgrade Training	
<i>By Major Brad Kelly (USAF), EdD</i>	
Chapter 8	190
Generative AI in the PME Classroom: Lessons Learned	
<i>By Kristan J. Wheaton</i>	
Part 3. Assessment	
Chapter 9	211
Enhancing Online PME Assessment: Design-Based Research for Rubric Development and Evaluation in the Global College of Professional Military Education	
<i>By Robert S. Nyland, PhD; and Carl P. Watts, PhD</i>	
Chapter 10	244
Experiments and Testing: Evidence-Based Instructional Strategies at the U.S. Naval Community College	
<i>By Charles B. Cushman Jr., PhD; Russell Evans; Daniel Goff; Scott Hamm, PhD; and Thomas Robertson</i>	
Chapter 11	271
Oral Comprehensive Exams: Lessons on the Assessment of Student Achievement of Program Learning Outcomes at the U.S. Army War College	
<i>By Colonel Christopher J. Hickey (USA), PhD; Professor Brett D. Weigle, EdD; and Professor Matthew C. Woessner, PhD</i>	

Chapter 12	299
Misconceptions among U.S. Air Force Academy Aviation Weather Cadets: An Exploratory and Validation Study Using a Survey of Meteorology Concepts	
<i>By Lieutenant Colonel Daniel O’Keefe (USAF), PhD; Major Louis J. Cascino (USA); Matthew Mynk; Senior Master Sergeant Nathaniel George (USAF); Major Christopher Francis (USAF); Wilson González-Espada, PhD; and Kimberly de La Harpe, PhD</i>	
Chapter 13	323
Data and Artificial Intelligence Literacy: Using Metrics and Experiential Learning to Inform Andragogy	
<i>By Kathleen Moore, PhD; and William Barry, PhD</i>	
Select Biography	345
Index	367
About the Authors	371



FOREWORD ●

The congressionally mandated Commission on the National Defense Strategy included the following, unsettling, top-line conclusion in its July 2024 report:

The threats the United States faces are the most serious and challenging the nation has faced since 1945 and include the potential for near-term major war. The United States last fought a global conflict during World War II, which ended nearly 80 years ago. The nation was last prepared for such a fight during the Cold War, which ended 35 years ago. It is not prepared today.¹

The commission blames this lack of preparation on the actions—and inaction—of several presidential administrations, a dysfunctional budgeting process in Congress, a withered defense industrial base, and a civilian population largely detached from and ignorant of national security concerns. The commission, however, notes that the all-volunteer force pro-

¹ Jane Harman et al., *Report of the Commission on the National Defense Strategy* (Washington, DC: Commission on the National Defense Strategy, 2024), v.

vides the United States with an “unmatched advantage.”² This echoes the confidence expressed for years in the quality of American servicemembers by senior leaders across the Department of Defense (DOD), members of Congress, the executive branch, and thought leaders across the country. Many of these same leaders—a succession of Service chiefs, chairmen of the Joint Chiefs of Staff, and secretaries of defense—consider professional military education (PME) a critical, indeed central, component in ensuring that American (and allied) military personnel continue to generate that unmatched advantage.

In December 2023, hundreds of military educators, including civilians and uniformed members, took part in the Fourth Annual Military Scholarship of Teaching and Learning (MSOTL) Forum held at Marine Corps University. The theme of this hybrid conference was “The Future of Military Education: Technology and the Art of War.” This conference and the series of which it is a part provide a venue for information-sharing and professional growth for PME, Service academy, and professional continuing education faculty from across the DOD and allied countries. This, in turn, directly supports faculty in fulfilling their central role of educating future generations of military, civilian, and allied leaders in national security.

Military education institutions face several challenges not generally shared by their civilian counterparts. I spent more than 18 of my 22 years as a civilian at Air University serving in administrative roles and personally grappled with these challenges every day. I remained active in the classroom at the same time and experienced these stressors from the perspective of a “line faculty member” as well. For example, whether officers or enlisted, PME students represent a wide variety of educational levels and backgrounds. In many cases, neither their previous academic work nor their professional experience is related to curricula that is focused on international security or Joint warfighting. This is particularly challenging in those PME programs taught at the graduate level. Knowing

² Harman et al., *Report of the Commission on the National Defense Strategy*, ix. The commission’s report does, however, note that this force is currently too small and “stressed” to prevail in a multitheater, global conflict.

how these adult professionals learn during their careers is as critical to an instructor's success as knowing the subject matter they are teaching.

At the same time, faculty in PME institutions, like their students, come from a variety of backgrounds and with often widely varying levels of experience as educators and military practitioners. This poses very real and enduring challenges for faculty recruitment, development, and retention, whether military or civilian. Administrators in military education programs typically face a level of faculty turnover—the natural result of the military assignment process—that would stagger their civilian counterparts. For example, military members constitute the large majority of faculty across all PME programs and rarely remain in place for more than three or four years, if that. College and school commandants likewise come and go quite frequently, a situation that all-too-often results in a lack of continuity and program changes driven by little more than the personal preferences of the “new boss.”

The possession of an advanced degree or the leadership experience gained through several supervisory, staff, or command assignments does not necessarily guarantee one's success in the development and delivery of a challenging and dynamic curriculum or in the accurate and meaningful assessment of their students' work. Successfully leading a seminar of experienced (and sometimes highly opinionated) officers and enlisted leaders while providing them with meaningful feedback on their work can be a challenge for even experienced educators. Recently, the DOD and the chairman of the Joint Chiefs of Staff directed the adoption of an outcomes-based versus the more traditional input-based approach across all military education.³ Accurately assessing student learning and the attainment of program-level outcomes is critical as the Services and Congress are tempted to eliminate programs that cannot prove a sound return on investment. Finally, PME programs' curricula are constantly changing to remain (rightly) responsive to the direction of senior civilian and

³ See *Department of Defense Instruction (DODI) 1322.35, Volume 1. Military Education: Program Management and Administration* (Washington, DC: Department of Defense, 26 April 2022); see also “Guidelines for Outcomes-Based Military Education Assessments,” in *Outcomes-Based Military Education Procedures for Officer Professional Military Education* (Washington, DC: Chairman of the Joint Chiefs of Staff, 2022), encl. D.

military leaders and relevant to an international security environment wracked by rapidly changing technology and geopolitical upheavals, requiring faculty to stretch into new fields of study and pushing them beyond their comfort zones.

The subjects analyzed and examined in this collection reflect the challenges noted above and more, but they also point to the value of sustaining the MSOTL Forum and other, similar opportunities. Administrators and faculty alike benefit greatly from the continuous exchange of ideas and the insights gained through applied research in the fields of teaching and learning. Far too frequently in my experience, however, harried academic leaders tended to focus on helping faculty, especially new faculty, with *what* to teach rather than *how* best to teach it. Moreover, individual student and overall program assessment was often viewed as someone else's job or written off as a bureaucratic drill. Nothing could be further from the truth.

This nation's armed forces and those of our allies clearly face a daunting international security environment. Maintaining our edge in human capital will be the best guarantee of success in deterring and, if necessary, defeating any opponent at the lowest possible cost in blood and treasure. Professional military education as well as professional continuing education play a central role in maintaining that edge. The MSOTL Forum, as the works contained in this volume clearly demonstrate, is likewise a critical resource in the professional development of those who will continue to develop and execute these vital programs.

Mark J. Conversino, PhD
Professor Emeritus, Air University



PREFACE ●

This year marks the fifth anniversary of the Military Scholarship of Teaching and Learning (MSOTL) Forum, an annual educational research conference focused on the military learning ecosystem. In those five years, my coeditor of this volume, Dr. Lauren Mackenzie, and I have witnessed some incredible things—research projects that start as good ideas and grow to published findings such as those in this volume, unlikely multinational partnerships that begin as professional networking and grow to personal friendships, and a groundswell of support for our objective of encouraging evidence-based teaching methods. However, one of the most unexpected developments to result from the MSOTL movement is linguistic in nature. As we hosted a recent MSOTL Forum onsite at Marine Corps University, the hallways were abuzz during breaks with discussions about the research presentations, yes, but also about MSOTL as a community, as a field of scholarly inquiry, as a noun, and as a verb. “How did you start MSOTLing?” and “MSOTL has changed how we think about teaching” were several overheard conversations.

A few weeks after that, Lauren and I fielded a call with colleagues at a North Atlantic Treaty Organization (NATO) ally military school who asked us, “How is your MSOTL center funded?” and “How does your MSOTL program operate?” Readers . . . there is no MSOTL center. There is no MSOTL program . . . at least not yet. Until now, MSOTL has been

Lauren and me working together in our free time on this labor of love because we know it is important and we know it can make a difference in the lives of our warfighters. It is clear now that many others across our global military learning environment feel the same!

Also at that recent MSOTL Forum, a new acquaintance asked me, “How did you first hear about MSOTL?” I knew, then, that Lauren and I had created something that would last. It was no longer about us having launched a conference; it had grown into an energized community of scholar-practitioners stretching from Kadena Air Base in Okinawa, Japan, to the Nigerian Defence Academy in Kaduna and everywhere in between. This edited volume is one culmination of that energy.

In *The Republic*, Plato wrote that “the beginning is the most important part of the work,” but I disagree. The most important part of MSOTL is that those who believe in it continue to carry it forward into their military classrooms around the world.

Megan J. Hennessey, PhD

Provost

Marine Corps University



GLOSSARY OF TERMS ●

ABM	air battle manager
AI	artificial intelligence
APIs	application programming interfaces
AR/VR	augmented reality/virtual reality
BMT	Basic Military Training (U.S. military)
CSR	case study research
DBR	design-based research
DEP	distance education program
DOD	Department of Defense
DON	Department of the Navy
<i>École de Guerre</i>	French War College
EPS	Education Participation Scale
GCPME	Global College of Professional Military Education
GenAI	generative artificial intelligence
GPT	generative pretrained transformer
ILE	immersive learning environments
IRB	institutional review board
JPME	Joint professional military education
LEI	learner empowerment instrument
LII	Leadership and Innovation Institute
LLMs	large language models

MilSOARR	Military Spouse Online Autism Relocation Readiness program
MLLMs	multimodal large language models
MSOTL	Military Scholarship of Teaching and Learning
Narrow AI	narrow-artificial intelligence
OBME	outcomes-based military education
PD	professional development
PLOs	program learning outcomes
PME	professional military education
QEP	Quality Enhancement Plan
SACSCOC	Southern Association of Colleges and Schools Commission on Colleges
S&W	simulation and wargaming
SEE model	student experience ecosystem model
STEM	science, technology, engineering, and mathematics
TAP	Transition Assistance Program (Department of Veterans Affairs)
TLCoP	teaching and learning community of practice

THE MILITARY SCHOLARSHIP OF TEACHING AND LEARNING



INTRODUCTION ●

Lauren Mackenzie, PhD

In their seminal 1995 publication, Robert B. Barr and John Tagg identified a shift in higher education moving toward the facilitation of learning in which teaching practice is focused, enhanced, and assessed.¹ It is in this spirit that the Military Scholarship of Teaching and Learning (MSOTL) Forum was originally convened and continues to grow. That is, regardless of an educator's academic background or the learning level at which they teach, the Scholarship of Teaching and Learning (SOTL) field offers meaningful methods for educators to put into practice and valid measures for researchers to integrate into their classroom assessments.

The chapters included in this edited volume highlight such practices and measures and, in doing so, create the kind of educational synergy that has been the hallmark of the MSOTL community since its first informal gathering in 2019. By way of background, the Marine Corps University Strategic Plan (Goal 5.4.4) tasks its Faculty Council to: “Engage other PME institutions in discussion for development of a society/association of Professional Military Educators as an opportunity to in-

¹ Robert B. Barr and John Tagg, “From Teaching to Learning: A New Paradigm for Undergraduate Education,” *Change: The Magazine of Higher Learning* 27, no. 6 (1995): 12–26, <https://doi.org/10.1080/00091383.1995.10544672>.

crease service school collaboration.”² In response to this call, a small group of military educators from across the United States and Canada came together in 2019 at the InterUniversity Seminar on Armed Forces & Society conference to collaborate and promote decision-making in military education through the lens of the scholarship of teaching and learning. Several months later in 2020, the first Joint Scholarship of Teaching & Learning Forum came to life, hosted by Dr. Megan Hennessey and her colleagues at the U.S. Army War College. This initiative expanded the traditional SOTL emphasis on/by civilian universities to create space for a military perspective in this growing field.

The chapters in this edited volume represent a further broadening of the relevance and applicability of the scholarship of teaching and learning for the military educator. Taken together, they represent a unique set of academic disciplines (ranging from meteorology to military history), methodologies (qualitative and quantitative), and learning levels (ranging from the community colleges to top-level schools). This diverse representation aligns with a recent characterization of SOTL as “transdisciplinary,” and we are delighted to include contributors from a variety of U.S. PME institutions, including the Army War College, Marine Corps University, the U.S. Air Force Academy, the U.S. Naval Academy, Air University, as well as international institutions such as the French War College.³ The MSOTL community that has formed during the past five years has served an important role in advancing research at various stages of development, and at the fourth annual (but first hybrid) MSOTL Forum in December 2023, these scholars presented the outcomes of their classroom inquiries at various stages, ranging from initial stages to completed research projects.

² *Marine Corps University Strategic Plan, 2022–2027* (Quantico, VA: Marine Corps University, 2022).

³ Laura Cruz, “Converging Spaces?: Re-Imagining the Scholarship of Teaching and Learning as the Subject of Transdisciplinary Research,” *Transformative Dialogues: Teaching and Learning Journal* 17, no. 1 (Spring 2024), <https://doi.org/10.26209/td2024voll7iss11823>; and Janice Miller-Young and Nancy L. Chick, eds., *Becoming a SoTL Scholar* (Elon, NC: Elon University Center for Engaged Learning, 2024), <https://doi.org/10.36284/celelon.oa6>.

In a recent article documenting the process of creating and assessing SOTL communities of practice, Laura A. Lukes and her research colleagues note that “SOTL requires intentional multi-level support to thrive at universities.”⁴ This message serves as an important reminder that the resources required to create a SOTL community are not the same as those required for it to thrive. That is, within the PME community, the MSOTL Forum has continued to grow as a result of its intentional prioritization by Dr. Hennessey and her team each year to bring the forum to life, and because it has consistently drawn from four levels of university support (provost-level, director-level, faculty/instructor-level, and academic support-level). For example, the 2023 MSOTL Forum (from which the chapters in this book are drawn) included welcome remarks from the MCU provost, the Air University Center for Teaching and Learning director as the co-lead for the event, faculty members who served as presenters and panel chairs, and Marine Corps educational officers from the MCU academic support division who worked as MCU ambassadors to support the virtual participants. In addition, the 2023 forum received more than 80 submissions, all of which were peer reviewed by individuals from across these levels, which led to 13 presentations that were converted into chapters (that were further peer reviewed) for the current edited volume.

Thus, in keeping with the theme of the forum—Creating Creative Learning Communities—as well as the objective of SOTL, the chapters included in this inaugural edited volume are organized as follows into three broad categories:

1. Community and creativity
2. Training and education case studies
3. Assessment

These 13 chapters highlight the continued innovation undertaken

⁴ Laura A. Lukes et al., “Impact of a Regional Community of Practice for Academic Developers Engaged in Institution-level Support for SoTL,” *International Journal for Academic Development* 29, no. 3 (2023): 393, <https://doi.org/10.1080/1360144X.2022.2135005>.

by professionals from across the continuum of military learning to excel both in the content they teach and the ways in which they teach it. It is an honor to amplify the research findings of scholars whose commitment to learning and growth serves as a model for future military educators.

Part 1

COMMUNITY & CREATIVITY



CHAPTER 1 ●

From Mindset and Collaboration to Connections and Experience

Creating a Teaching and Learning Community of Practice
in the U.S. Air Force Leadership and Innovation Institute

*John Hinck, PhD; Steven Davis, PhD; Robert Hinck, PhD;
Sara Kitsch, PhD; Mary Bartlett, PhD; and Andy Clayton, EdD*

In 2017, Air University (AU) created the Leadership and Innovation Institute (LII) in Montgomery, Alabama, as a dedicated center to coordinate leadership education at AU and the U.S. Air Force writ large. Its mission was not to force curriculum onto other programs or to centralize or standardize the execution of leadership curriculum. Instead, LII was intended to build a sense of community around leadership education with a dedicated staff of leadership subject matter experts. This mission was both aspirational and vague, and within its first two years of existence, LII took on two additional duties: 1) executing the planning and assessment of AU's five-year Quality Enhancement Plan (QEP); and 2) supplying the teaching, training, and curricular expertise for the Leader Development Course for Squadron Command (LDC). Despite its limited staff, growing mission set, and little organizational authority to enforce change, LII proved largely successful in delivering impactful leadership and leadership development curriculum due to its creation of a teaching and learning community of practice (TLCoP), with unique emphasis on the importance of mindset, collaboration, connection, and experience.

This chapter examines LII as a case study of approaching professional military education (PME) through TLCoP and demonstrates the benefits associated with doing so. It argues that LII—with a loosely connected staff with diverse academic backgrounds, professional experiences, and

variated teaching obligations—innovatively impacted leadership teaching and learning curriculum across AU and the Air Force due to LII’s novel approach to faculty development, incorporation of immersive learning technologies, and engaging content in areas of ethics, empathy, and post-vention. These efforts proved impactful, evidenced by student and faculty feedback and incorporation of LII’s teaching and learning curriculum by programs outside of LII, and occurred due to LII’s formation of a TLCoP rather than the use of traditional teaching and learning models.

Overview of the Leadership and Innovation Institute

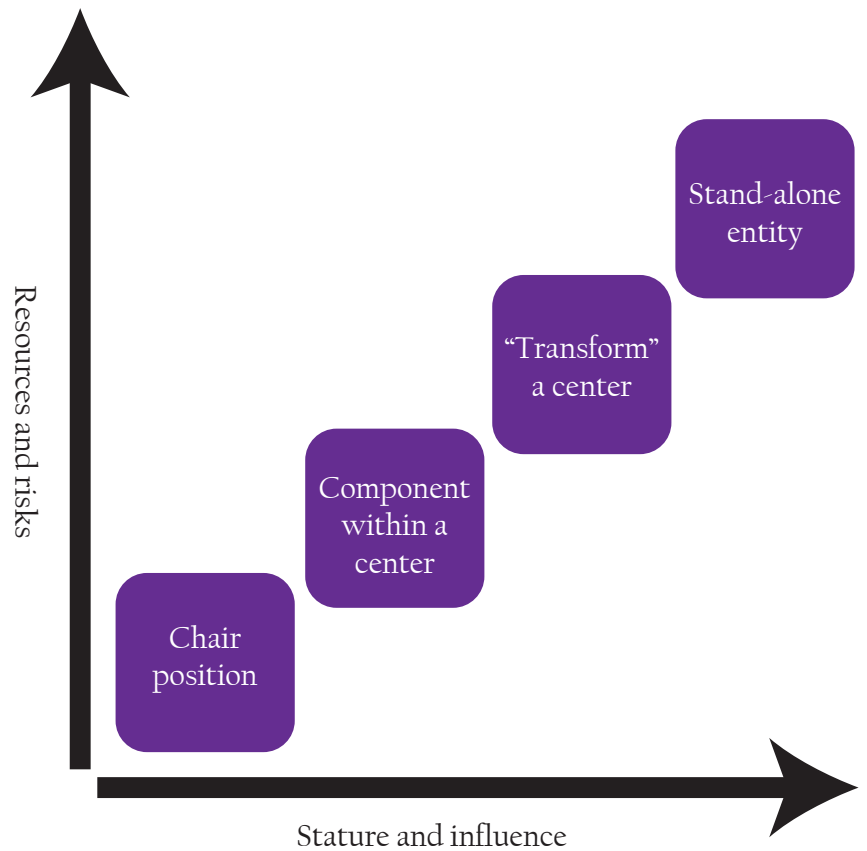
The mission of LII is to deliver a continuum of leadership education, research, and outreach for the Air Force and AU.¹ LII emerged from discussions at AU in 2015 exploring potential roles, resources, and organizational structures for an “AU Leadership Center.” Options explored included various levels of resources relative to potential formal organizational stature and influence within AU (figure 1). These discussions concluded that AU *did* need a leadership center to serve as a dedicated hub for leadership research and outreach but that such a center *would not* function as a centralized arm for executing leadership curriculum across AU; that it *would not* serve to standardize curriculum across AU schools/programs; or would it act like an inspector general toward other schools’ leadership curriculum. Hence, in contrast to other PME universities centers for leadership and ethics, LII’s mission and teaching duties are largely similar but its formal organizational role and authority is markedly lower.²

LII is comprised of 11 civilian employees who include a director and 10 administratively determined (AD) faculty, the majority of whom hold

¹ See “Leadership and Innovation Institute,” Air University, accessed 19 September 2024.

² For example, Naval War College has the College of Leadership and Ethics, Army War College has the Center for Strategic Leadership, and Marine Corp University has the Lejeune Leadership Institute. All three organizations share similar mission sets as LII with a focus on teaching formal leadership curriculum across educational programs and providing outreach for broader interested parties. Although focusing on undergraduate education, the Air Force Academy’s Center for Leadership and Character Development reflects a far more similar center for leadership compared to other PME institutions relative to AU’s LII.

Figure 1. Potential courses of action (COAs) for establishing an AU leadership center



Source: courtesy of the authors, adapted by MCUP.

doctorate degrees in fields ranging from history, communication, leadership, and counseling education. LII is administratively assigned to Air War College (AWC) through a memorandum of agreement between the Ira C. Eaker Center for Professional Development and AWC. Much of LII’s leader development mission occurs at Eaker’s LDC, while LII faculty also teach core curriculum and electives at AWC and Air Command and Staff College (ACSC), in addition to providing outreach at nearly every

school/center across AU. Taken together, LII faculty are charged with three lines of effort.

Education

- *Leader Development Course (LDC)*: LII faculty teach, create curriculum, and train the incoming military faculty (graduated squadron commanders) who arrive each summer and depart for AWC the following summer. While this course is executed by the Commander's School, LII faculty provide the educational expertise, faculty development, and organizational continuity ensuring the course's continued success.
- *Project Mercury*: In collaboration with the University of Michigan in Ann Arbor and the Innovatrium, LII executes this 12-week course that provides innovation education curriculum to students Air Force wide. Students are placed in intentionally diverse teams that grapple with challenging problem sets while being guided by experienced coaches and alumni mentors. Students graduate with the tools they need to build the innovation culture, competency, and community the nation needs.
- *Leadership Coaching Program*: LII's coaching program, approved for board-certified coach (BCC) credentials, is an eight-day course (60 hours) of self-study/preparation in person and virtual (Zoom) experiences that includes practice with peers and two live assessments.
- *Additional Teaching across AU*: LII is expected to provide one to two instructors annually to teach ACSC's Leadership and the Profession of Arms Course. LII faculty are further expected to create and teach electives for AWC and ACSC.

Research

- *Quality Enhancement Plan (QEP)*: AU's five-year QEP is required for reaffirmation of accreditation by the Southern Association of Colleges and Schools Commission on Colleges (SACSCOC). AU's QEP topic focuses on "Leadership and Ethics across the Continu-

um of Learning.”³ Informed by an intensive leadership and ethics literature review, AU developed a conceptual ethical leadership framework and competencies. After curriculum mapping, focus groups, and surveys, AU focused on three ethical leader focus areas: empathy, ethical decision making, and fostering innovation. Targeting select schools, the QEP team (the QEP director is dual hatted as the LII director) has worked with select courses within the schools to assess and adjust curriculum to improve student outcomes. AU is currently in year five of the QEP, with a report due to SACSCOC by March 2025.

- *Individual Faculty Research:* Aside from LII’s intensive research associated with the QEP, LII faculty are expected to conduct research contributing to their academic disciplines. LII publications and presentations include suicide postvention, measuring the impact of LDC, leadership coaching as a transformative experience, assessing developmental stages and faculty development, and assessment of Chinese and Russian strategic narratives, to name a few.⁴
- *Focused Faculty Research:* LII is currently working with AU/Operations Directorate (A3) on the Comprehensive Airmen Fitness research initiative and have members serving on the Headquarters Air Force/Manpower, Personnel, and Services’ (A1) Leadership Assessment Action Group and the Human Performance Steering Group. LII has also created an LII research plan with teams researching the history and practices of leadership development within the Air Force and other U.S. military Services, theories of leadership and leader development within academia, and various measures of effective leadership and leadership training.

³ *Quality Enhancement Plan (QEP): Leadership and Ethics Across the Continuum of Learning* (Birmingham, AL: Air University Press, 2019).

⁴ Examples of faculty research can be found at “LII Research,” Air University, accessed 20 September 2024.

Outreach

- *Workshops and Lessons*: LII faculty are in high demand to present human domain and soft skills not only to AU schools and centers but also audiences Air Force wide. LII offers workshops in empathy and ethical decision making and lessons on energy management, personality typology, emotional intelligence, design thinking, innovation, small group communication, resiliency, and suicide pre/postvention. In FY 2023, LII faculty reached 10,298 students through outreach efforts including in-person and virtual offerings of various workshops that amounted to 864 hours of leadership material.
- *Virtual LeaderShops*: Because of high demand for leadership development opportunities, LII began offering a virtual, monthly leadership workshop (LeaderShop). Every first Friday of the month, LII facilitates a lesson on various leadership topics, advertised through the Teaching and Learning Center and AF Portal. LII averages around 300 registrants per LeaderShop.

The scope and range of LII's activities, in combination with its members' diverse research backgrounds and interorganizational positioning within multiple AU programs and schools, creates inherent tensions and challenges. Similar to critiques of organizational stovepiping within U.S. government bureaucracies, including PME, academics at civilian higher-education institutions are noted for retreating into academic silos while strategically reducing their commitments to organizational service due to its lack of consideration into promotion decisions.⁵ And yet, faculty at LII appear to have effectively navigated these tensions—evident by the range of courses offered, substantial outreach conducted, and accolades rewarded. How LII manages to create and deliver this range of curriculum, with its increased spread and demand from other programs, can unveil useful models for other PME organizations.

⁵ Kerry Ann O'Meara, "Ensuring Equity Service Work," *Inside Higher Ed*, 9 May 2018.

Teaching and Learning Community of Practice

The concept of teaching and learning communities emerged in the early 1990s when educational researchers and practitioners identified learning as occurring within shared practices between students and teachers.⁶ As a departure from the traditional view of schooling whereby individual knowledge and performance is emphasized, with students assumed to acquire the same body of knowledge at the same time, the defining quality of a teaching and learning community is a culture of learning whereby everyone is involved in a collective effort of understanding.⁷ Hence, rather than viewing education as process by which instructors pass information to students, learning becomes a collaborative process with input and shared authority among teachers and students collectively sharing a common learning goal. Culturally, teaching and learning communities share four characteristics: a diversity of expertise, a shared objective, an emphasis on learning how to learn, and a mechanism for sharing what is learned.⁸

One of the central themes to creating a community of teaching and learning for both faculty and students is experienced-based learning. Whether the focus is on a specific lesson, a course, or faculty development, creating visceral experiential learning moments for faculty and students is the glue that allows for the development of leadership behaviors for students and faculty as well as educational development for faculty, staff, instructors, and learners. Research into TLCoP identifies 16 areas influencing the creation and sustainment of effective teaching and learning communities, ranging from the role of teachers and learners (participants as both learner and teacher, coteaching, environment, role of school leaders); instructional techniques and learning methodologies (constructive

⁶ Wolff-Michael Roth and Yew-Jin Lee, "Contradictions in Theorizing and Implementing Communities in Education," *Educational Research Review* 1, no. 1 (2006): 27–40, <https://doi.org/10.1016/j.edurev.2006.01.002>.

⁷ Marlene Scardamalia and Carl Bereiter, "Computer Support for Knowledge-Building Communities," *Journal of the Learning Sciences* 3, no. 3 (1994): 265–83, https://doi.org/10.1207/s15327809jls0303_3.

⁸ Katerine Bielaczyc and Allan Collins, "Learning Communities in Classrooms: A Reconceptualization of Educational Practice," in *Instructional Design Theories and Models*, vol. 2, *A New Paradigm of Instructional Theory*, ed., Charles M. Reigeluth (Mahwah, NJ: Lawrence Erlbaum Associates, 1999), 16–18.

conflict, experiential, immersion, use of technology, intervention—action versus remediation); relational dynamics (relationships, collaboration, commitment, transparency climate); and commitment to improvement (focus on learning/outcomes, dynamic cycles of learning, innovation).⁹

⁹ Jacalyn M. Flom, Karen M. Green, and Steven Wallace, “Helping Higher Education Students Succeed: An Examination of Student Attributes and Academic Grade Performance,” *Active Learning in Higher Education* 24, no. 2 (2013): 221–35, <https://doi.org/10.1177/14697874211055689>; Kristin N. Rainville, Cynthia G. Desrochers, and David G. Title, eds., *Faculty Learning Communities: Communities of Practice that Support, Inspire, Engage and Transform Higher Education Classrooms* (Charlotte, NC: IAP, 2014); Markku Antinluoma, Liisa Ilomäki, and Auli Toom, “Practices of Professional Learning Communities,” *Frontiers in Education* 6 (2021): 1–14, <https://doi.org/10.3389/feduc.2021.617613>; Christopher M. Clark et al., “Community of Practice Among Faculty Team-Teaching Education Doctorate (Ed.D.) Students: A Reflective Study,” *International Journal of Doctoral Studies* 16 (2021): 379–93, <https://doi.org/10.28945/4775>; Lorin W. Anderson, “The Classroom Environment Study: Teaching for Learning,” *Comparative Education Review* 31, no. 1 (1987): 69–87, <https://doi.org/10.1086/446656>; Sue Bennett and Lori Lockyer, “Becoming an Online Teacher: Adapting to a Changed Environment for Teaching and Learning in Higher Education,” *Educational Media International* 41, no. 3 (2004): 232–36, <https://doi.org/10.1080/09523980410001680842>; John M. Hinck et al., “Understanding ‘Leadergogy’: The Unique Teaching and Learning Methods in a Leader Development Course,” *Journal of Military Learning* (October 2023): 22–45; David Kember, Amber Ho, and Celina Hong, “The Importance of Establishing Relevance in Motivating Student Learning,” *Active Learning in Higher Education* 9, no. 3 (2008): 249–63, <https://doi.org/10.1177/1469787408095849>; Adolf Winteler, “The Academic Department as an Environment for Teaching and Learning,” *Higher Education* 10 (1981): 25–35; Antinluoma, Ilomäki, and Toom, “Practices of Professional Learning Communities”; Benjamin D. Brown, Rolanda S. Horn, and Gwendolyn King, “The Effective Implementation of Professional Learning Communities,” *Alabama Journal of Educational Leadership* 5 (August 2018): 53–59; Richard DuFour, “What Is a Professional Learning Community?,” *Educational Leadership* 61 (2004): 6–11; David A. Kolb, *Experiential Learning: Experience as the Source of Learning and Development*, 2d ed. (Upper Saddle River, NJ: Pearson, 2015); Alice Y. Kolb and David A. Kolb, “Experiential Learning Theory: A Dynamic, Holistic Approach to Management Learning, Education and Development,” in *The SAGE Handbook of Management Learning, Education and Development* (London: Sage, 2009), 42–68, <https://doi.org/10.4135/9780857021038.n3>; Charles M. Reigeluth and Yunjo An, “What’s the Difference between Learning Experience Design and Instructional Design?,” *Journal of Applied Instructional Design* 12, no. 3 (January 2023): 227–40, <https://doi.org/10.59668/515.12897>; Joseph T. Wong and Bradley S. Hughes, “Leveraging Learning Experience Design: Digital Media Approaches to Influence Motivational Traits that Support Student Learning Behaviors in Undergraduate Online Courses,” *Journal of Computing in Higher Education* 35 (2023): 595–632, <https://doi.org/10.1007/s12528-022-09342-1>; Paul Tosey, “The Learning Community: A Strategic Dimension of Teaching and Learning?,” in Paul Jarvis, ed., *The Theory and Practice of Teaching*, 2d ed. (New York: Routledge, 2006), 130–46, <https://doi.org/10.4324/9780203016442>; Dennis E. Beck, Leonel Morgado, and Patrick M. O’Shea,

Method

Research Design

As a methodology, the case study research (CSR) approach is a qualitative method used for shaping the research design based on the appropriateness of the research questions, as well as to connect the purpose or question with the outcome or answer.¹⁰ CSR involves a holistic approach incorporating multiple levels of constructivism that is a bounded three-reality view involving the external reality that is brought to the

“Finding the Gaps about Uses of Immersive Learning Environments: A Survey of Surveys,” *Journal of Universal Computer Science* 26, no. 8 (2020): 1043–73, <https://doi.org/10.3897/jucs.2020.055>; Sara De Freitas et al., “Learning as Immersive Experiences: Using the Four-Dimensional Framework for Designing and Evaluating Immersive Learning Experiences in a Virtual World,” *British Journal of Educational Technology* 41, no. 1 (2010): 69–85, <https://doi.org/10.1111/j.1467-8535.2009.01024.x>; Harm H. Tillema, “Embedding and Immersion as Key Strategies in Learning to Teach,” in Henny Boshuizen, Rainer Bromme and H. Gruber, eds., *Professional Learning: Gaps and Transitions on the Way from Novice to Expert* (New York: Springer, 2004), 141–56, https://doi.org/10.1007/1-4020-2094-5_8; Liz Barnett et al., *Using Technology in Teaching and Learning*, 1st ed. (London: Routledge, 1997); Kaite J. Carstens et al., “Effects of Technology on Student Learning,” *Turkish Online Journal of Educational Technology* 20, no. 1 (2021): 105–13; David Pratt, Neville Davies and Doreen Connor, “The Role of Technology in Teaching and Learning Statistics,” in Carmen Batanero, G. Burrill, and C. Reading, eds., *Teaching Statistics in School Mathematics: Challenges for Teaching and Teacher Education* (Dordrecht, Netherlands: Springer, 2011), https://doi.org/10.1007/978-94-007-1131-0_13; Clarke et al., “Community of Practice”; Nancy A. Mundschenk and Wendy W. Fuchs, “Professional Learning Communities: An Effective Mechanism for the Successful Implementation and Sustainability of Response to Intervention,” *Southeastern Regional Association of Teacher Educators Journal* 25, no. 2 (2016): 55–64; Antinluoma, Ilomaki, and Toom, “Practices of Professional Learning Communities”; Ferlazzo, “A Professional Learning Community”; Antinluoma, Ilomaki, and Toom, “Practices of Professional Learning Communities,” 12; Brown, Horn, and King, “The Effective Implementation of Professional Learning Communities,” 53–59; DuFour, “What Is a Professional Learning Community,” 9; Larry Ferlazzo, “A Professional Learning Community Is Not a Faculty, Grade Level, or Department Meeting,” *Education Week*, 27 April 2021; Antinluoma, Ilomaki, and Toom, “Practices of Professional Learning Communities,” 1–14; DuFour, “What Is a Professional Learning Community,” 6–11; Micheline T. H. Chi and Ruth Wylie, “The ICAP Framework: Linking Cognitive Engagement to Active Learning Outcomes,” *Educational Psychologist* 49, no. 4 (2014): 219–43, <https://doi.org/10.1080/00461520.2014.965823>; Clarke et al., “Community of Practice,” 55–64; Brown, Horn, and King, “The Effective Implementation of Professional Learning Communities”; and Clarke et al., “Community of Practice,” 55–64.

¹⁰ Robert E. Stake, *The Art of Case Study Research* (Thousand Oaks, CA: Sage, 1995), 1–14; and Robert K. Yin, *Case Study Research: Design and Methods*, 5th ed. (Thousand Oaks, CA: Sage, 2014), 25–65.

situation, an experiential reality about how to engage and interpret based on actions between people, and a rational reality that integrates the collective interpretations.¹¹ CSR is heuristic, providing a better understanding of the phenomenon under study; is more contextual, conceptual, and based on experiential knowledge, and can capture the social construction of multiple realities.¹² Furthermore, CSR advocates a practical approach to collecting available information and using the knowledge to affect future decisions, as well as “being logical, analytical, and critical in fitting individual values with program development, constructing a program from several realities, thinking creatively about concepts and programs, and being able to shift attention between perspectives in the ongoing research process.”¹³

To support the findings with triangulation, the CSR approach discusses the findings based on guiding theory, multiple levels of evidence and interpretation, participant voices, and survey results.¹⁴ In response to answering the research questions within the context, CSR is the best fit and the most methodological sense under the circumstances. Additionally, CSR is beneficial “using theoretical propositions to guide data collection and analysis, along with allowing the researcher to better address the contextual, concrete, and experiential research aspects involving increasing human and organizational potential and to better capture the complexities of leader development as a systematic approach for individual and organizational growth.”¹⁵ CSR goes beyond the bounded system and offers a “thick” description of what is learned from data interpretation based on the participants’ cognitive understanding and affective experiences.

¹¹ Stake, *The Art of Case Study*, 71–90; Corrine Glesne, *Becoming Qualitative Researchers: An Introduction*, 3d ed. (Upper Saddle River, NJ: Pearson, 2016), 13; and John M. Hinck, “Designing an Organizational Leadership Development Program: A Case Study Approach,” *Journal of Nonprofit Education and Leadership* 7, no. 4 (2017): 268–86, <https://doi.org/10.18666/JNEL-2017-V7-I4-8081>.

¹² Sharan B. Merriam and Elizabeth J. Tisdell, *Qualitative Research: A Guide to Design and Implementation* (San Francisco, CA: John Wiley and Sons, 2009), 8.

¹³ Anne Sigismund Huff, *Designing Research for Publication* (Thousand Oaks, CA: 2009), 85–106; and Hinck, “Designing an Organizational Leadership Development Program,” 270.

¹⁴ Stake, *The Art of Case Study*, 108–10.

¹⁵ Hinck, “Designing an Organizational Leadership Development Program,” 270–71.

LII is studied as a case within the larger case of teaching and learning communities using a four-phase process:

- Phase 1. Literature review that informs the concepts of a TLCoP.
- Phase 2. Document analysis and organizational analysis to identify the frequency of teaching and learning concepts present in LII's teaching and learning efforts.
- Phase 3. Identifying the key components of LII's teaching and learning approach, conducted in informal discussions with LII faculty during a six-month period.
- Phase 4. Cross analysis of findings from phases 2 and 3.

Participants

The 11 faculty that comprise LII were the main participants in the case study, as they are the ones who embody the LII TLCoP. The indirect participants included the students and fellow instructors who completed end-of-course surveys in the programs and courses taught by LII faculty.

Data Collection

A meta approach was used in collecting data that had the potential to help understand LII's TLCoP, which comprised:

- 14 organizational documents (e.g., Air Force instruction manuals, mission statements, etc.).
- 29 organizational events with their inputs and outputs from LII meetings, faculty huddles, organizational and course improvement events, etc.
- 44 total established programs/courses (content and curriculum) taught by LII faculty (e.g., leadership development course, Project Mercury, core courses, electives, etc.) and internally created and delivered LII programs (e.g., leadership coaching program, faculty development, workshops, and LeaderShops).
- 5,418 previously collected and available end of course sur-

veys of all courses taught by LII faculty (e.g., core courses, electives, leader development courses, augmented reality/virtual reality (AR/VR) research task forces, postvention training, LII LeaderShops, etc.). The various course/program directors provided the end-of-course survey data.

- 117 scholarships published by LII faculty (e.g., books, book chapters, journal articles, scholarships of discovery, application, integration, etc.).
- 10 informal discussions among LII faculty throughout the research process.

Data Analysis

Data was analyzed using a combined approach of organizational analysis and document analysis. The cumulative coding began with an initial set of primary codes that were the 16 concepts from the literature review. The second coding cycle identified emergent codes. The third coding cycle was largely dependent on the LII faculty's analytical discussion on how they would rank the concepts based on quantity of appearance and based on quality of concepts. The final coding cycle involved cross analysis of findings to identify and categorize the concepts into levels representing LII's TLCoP.

Findings

The findings are presented in three ways: 1) phase 2 and 3 results presented in a table that maps to what extent the 16 concepts from literature and four emergent concepts “showed up” or were represented in LII's teaching and learning; 2) selected phase 2 results in a figure that depicts the relationship of connection that was developed based on how the data emerged as participants described the importance of connection with self, others, and the other; and 3) phase 4 results in a figure showing the ranking and organizing of the 20 concepts that form the primary, secondary, and tertiary levels of LII's TLCoP.

The three domains of connection represent the results of cross-case analysis of how participants (students *and* instructors) shared their con-

Table 1. Mapping the 20 concepts of teaching and learning to LII’s teaching and learning

	Organizational documents	Organizational events	Programs and events	End of course surveys	LII scholarship	LII teaching approach
	14	29	44	5,418	117	10
16 concepts from literature						
Collaboara-tion	10	21	6	98	4	10
Commitment	3	5	2	22	2	4
Constructive conflict	2	17	3	17	5	2
Coteaching	5	10	40	2,352	8	6
Dynamic cycles of learning	0	9	33	35	2	3
Environment	3	22	39	5,418	8	8
Experiential	4	27	44	5,418	16	7
Focus on learning and outcomes	4	22	44	5,418	10	8
Immersion	4	4	40	3,499	19	
Innovation	10	14	11	307	35	7
Intervention	0	8	0	14	3	2
Participants as learner and teacher	2	5	21	3,111	34	9
Relationships	2	24	38	2,201	55	6
Role of school leaders	1	2	0	67	4	3
Transparency climate	1	16	39	9	2	2
Use of tech-nology	3	19	44	3,190	13	5
Four concepts that emerged						
Experience	4	27	40	2,429	18	10
Mindset	5	8	12	1,151	9	10
Moments	1	26	38	4,914	15	10
Connection	3	28	44	5,002	71	10

Source: courtesy of the authors.

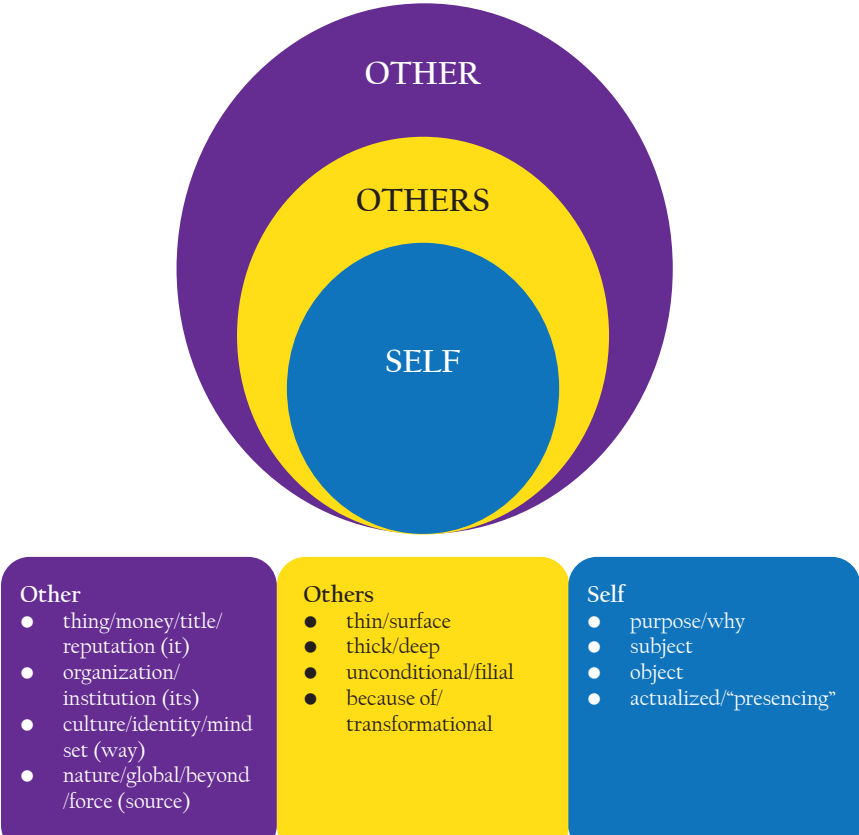
nection experiences with the informal discussions among the 10 LII faculty. The subthemes that emerged included:

1. While participants yearned and wanted to feel like they mattered and be connected to something bigger than themselves, they simultaneously discussed being afraid of connecting to self/others.¹⁶
2. Despite being “connected” through technology, participants felt alone in their humanity.
3. The depth of a lack of connection or “connection chasm” was present regardless of age, organization, and career field.
4. One of the top reasons why participants discussed leaving or quitting was because they felt like few people saw them, their identity, or due to a lack of connection.
5. Feeling connected to others was the highest component for what participants thought would be a good/strong organizational culture.

The emergent themes were conceptualized into three domains of connection (figure 2). The “self” domain centers on the connections someone has with oneself and the content that includes purpose/why, subject, object, and actualized “presencing.” The “others” domain is about the connections someone has with others in the learning environment that includes thin/surface, thick/deep, unconditional/filial, and because of/transformational. The other domain relates to connections that are more distant and objective, including the thing/money/title/reputation (it), the organization/institution (its), the culture/identity/mindset (way), and nature/global/beyond/force (source).

¹⁶ The data was collected from a study that was approved (or granted exemption) by the appropriate institutional and/or national research ethics committee (USAF IRB FWR20190084N with determination of NRNHSR). Additionally, the discussions among the research team were captured for some of the data analysis. The study was performed in accordance with the human research ethical standards as laid down in the 1964 Declaration of Helsinki and its later amendments or comparable ethical standards.

Figure 2. The three domains of connection



Source: courtesy of the authors, adapted by MCUP.

LII Teaching and Learning Community of Practice

The cross analysis from phases 2 and 3 revealed that the LII TLCoP contains all 20 teaching and learning concepts (16 from the literature and 4 emergent) and that a TLCoP is comprised of many concepts that can be organized into three levels (table 2). The six concepts in the primary level were ranked high among 9 or 10 of the faculty and showed up with high numbers with greater consistency across multiple categories. The eight concepts in the secondary level were consistent among six to eight of the

Table 2. Three levels of LII’s teaching and learning community of practice

Primary	Experience Mindset Moments Connections Collaboartion Participant as learner and teacher (shared authority)
Secondary	Coteaching Environment Experiential Focus on learning objectives Immersion Innovation Relationships Use of technology
Tertiary	Commitment Constructive conflict Dynamic cycles of learning Intervention Role of school leaders Transparency climates

Source: courtesy of the authors.

faculty and showed up very high in scholarship and in end of course surveys. The six concepts in the tertiary level were consistent among two to five of the faculty and, overall, showed up less across categories.

Discussion

The discussion section defines the four emergent concepts of experience, mindset, moments, and connection followed by a deeper discussion that applies the four new constructs in the broader principles of the educational and development aspects of LII’s TLCoP in three unique ways: the importance of experience in teaching and learning, how experience and mindset are key in developing faculty, and the combined use of experience, mindset, moments, and connection for ethics, empathy, and post-vention training.

Defining the Four Emergent Terms that Comprise LII's TLCoP Primary Level

Of the six concepts that comprise the primary level of LII's TLCoP, the four new ones are defined as specific to this research and could be generalized to the wider academic community. *Experience* refers to how the learning process comes alive for participants and culminates holistically that integrates the head (cognitive), heart (affective), hand (behavioral), and connections (relational) ways of immersion. *Mindset* refers to the frame of mind of participants before, during, and after learning. *Moments* refers the fact that most people and organizations often do not know a moment is a moment or how something is created until the experience and process can be looked back on and understood as a meaningful experience. *Connection* refers to the integrated relationships of the three domains of connection: self, others, and the other.

The Importance of Experience in Teaching and Learning

LII builds its teaching and learning communities of practice through the use of immersive learning and immersive learning environments (ILEs), which enables students to learn problem-solving, communication, emotional intelligence, situational awareness, and stress inoculation to related subject matters through simulated first-person ILEs. In various programs, LII uses a mixed reality avatar conversation simulation as its immersive learning simulation (ILS), which includes a human-in-the-loop system, where a trained simulation specialist (usually a specialized trained actor) controls the movement, interaction, and voice of an avatar in a mixed reality environment.¹⁷ More specifically, this type of human-in-the-loop simulation is referred to as an avatar-mediated interactive training and individual experience system (AMITIES).¹⁸ LII's hybrid approach of using both enabled artificial intelligence with live inhibitors from simulation specialists can be seen as a mixed reality leadership experience (MRLx),

¹⁷ Andrew S. Clayton and Carrie Straub, "Virtual Leadership Simulator: The Missing Gap in Soft Skills Training" (conference presentation, Interservice Industry Training, Simulation and Education Conference [I/ITSEC], Orlando, FL, 30 November–4 December 2020).

¹⁸ Clayton and Straub, "Virtual Leadership Simulator."

a mixed reality exercise (MRx), an immersive learning experience (ILx), and ILS.¹⁹

Through live interactive conversation experiences, learners communicate with avatars controlled by trained simulation specialists, creating synchronous, live, and authentic dialogue between two individuals. Students, after being given a scenario to start the simulation, determine how best to communicate with the character (avatar) to achieve the most effective results in a one-to-one dialogue. Because the conversation is live, learners practice a wide range of human domain skills to properly interact with the avatar based on the context of the scenario given to the learner.

LII's use of ILS is an exemplary example of utilizing technology to build learning communities where the learner is placed into an immersive learning experience that is learner-centered and directed. Skillsets such as ethical decision making, empathy, emotional intelligence, communication, and active listening—all key elements of human domain—are applied. After the simulation is complete, this immersive learning moves from the experience phase to the reflective or introspective phase, where the learner, along with other learners, observe the simulation to reflect on the content, the process, and the value of the experience and ultimately construct new knowledge about leadership and the practice of humanity. Throughout the experience, the learner is placed within an authentic and realistic experience of the human domain, allowing learners the opportunity to demonstrate and develop authentic leadership behaviors in a safe repeatable immersive learning environment.

As an illustration, one of the signature scenarios developed within the ILS simulator is titled, “Deployed Spouse” (figure 3).²⁰ In this scenario, the learner is presented with a short paragraph prior to the start of the simulation. The “Deployed Spouse” scenario handout paragraph is provided below.

Stacy Lewis is a military spouse whose husband, Staff Sergeant (SSgt) Tim Lewis, is currently deployed to Af-

¹⁹ Clayton and Straub, “Virtual Leadership Simulator.”

²⁰ “Day 8–Capstone Script and IT Checklist,” Leader Development Course for Squadron Command, 2024.

Figure 3. Learners interacting in the immersive learning simulation (ILS)



Source: Air University, courtesy of the authors.

ghanistan. A few weeks before SSgt Lewis deployed, SSgt Lewis and his wife Stacy had a baby girl. Now SSgt Lewis is deployed, and Stacy is having difficulty dealing with a deployed husband and raising their new daughter all by herself. Stacy has stopped by your office to speak to you about Tim being deployed and the struggles she is having as a single parent.

The objectives for this “Deployed Spouse” scenario require the learner to apply leadership and command decision-making behaviors to effectively communicate and navigate the issues the deployed spouse is experiencing. Another objective is for the learner to apply emotional intelligence during this scenario. The last objective is for the learner to appreciate the challenges faced with command. The objectives for this scenario are somewhat ambiguous. One might ask, “What behaviors of leadership should be displayed, or which aspects of emotional intelligence should

be demonstrated?” The answer to these questions is also ambiguous—it depends. Depending on the learner and the learner’s experience, leadership style, communication approach, personality, and level of emotional intelligence are just a few factors that will determine how the learner will reach these objectives. Additionally, the responses from the virtual avatar will also affect how the learner approaches the scenario. Figure 3 depicts a learner executing the ILS simulator and interacting in the “Deployed Spouse” scenario. A brief look at the dialogue between the learner and the virtual avatar (Stacy Lewis) provides a richer understanding of what is occurring in the ILS simulator.

[Learner] *“Start simulation”*

[Avatar] *“Hello, sir, thank you for letting me come to your office to talk to you.”*

[Learner] *“Of course, how are you doing? How can I help you?”*

[Avatar] (using a soft and quivering voice) *“Well, I wasn’t sure where else to go, but I was wondering if there is any way that my husband Tim can come home early from his deployment?”*

[Learner] *“Stacy, it sounds like there is something going on. Can you tell me more about how things are going at home?”*

[Avatar] (subtly crying and wiping tears from her face) *“Um, I am struggling with Tim being deployed and I don’t think I am cut out to be a single mom. That’s why I need Tim to come home. Is there any way he can come home?”*

[Learner] *“Stacy, I am so sorry you are feeling this way. It can be very difficult to not have your spouse with you with a new baby, I know, because when I was deployed and had to leave my wife to take care of our kids and I know she struggled at times too.”*

[Avatar] *“So you know what it is like, and you know that I really need Tim to come home. Can he come home?”*

[Learner] *“Stacy, Tim is doing a great job down range and an important job supporting the wing’s mission. I think maybe if we look at some other resources, we might be able to find some other options that can help you while Tim is deployed. Can we look at*

some of these other options like the spouses' program, the chaplain, or Military and Family Life Counseling (MFLC) program?"

[Avatar] "I just don't have the energy to do anything. A few spouses came by my house, but I did not let them in because my house is a mess. None of the chores are done and I wasn't even showered or dressed. I can barely get out of bed, and every time Emma cries, I feel like I have failed, and I don't feel connected to her. I feel like a ghost and only Tim can make this go away."

This is just a small part of the dialogue that the learner will encounter during the simulation. The learner who has troops deployed down range is now dealing with one of their troop's spouses who is struggling being a single parent and is asking—almost insisting—that her husband needs to come home from his deployment. How does the learner address this request? What happens if the learner says yes to Stacy? What impact does that have on Tim? What impact does that have on the down range unit? Should the learner have talked to anyone else before saying yes or no to Stacy? How did the learner handle the request from Stacy? Is bringing Stacy's husband home from his deployment going to fix the situation? What is the real issue that is going on? Did the learner recognize that Stacy may have postpartum depression? If so, how did the learner address that? How did the learner address what other resources may help Stacy? These are just a few of the thoughts and questions that the learner is asking themselves during the simulation to navigate the conversation and apply the behaviors outlined in the objectives for this scenario such as applying emotional intelligence, applying command decisions as the commander of the unit, and appreciating the challenges one will face as a commander.

LII's effective usage of immersive learning resulted in multiple programs incorporating similar learning activities in addition to collaborations among faculty and students at multiple AU schools. Dr. Andy Clayton heads these efforts at LII and leads a year-long, research intensive task force on immersive learning with students at AWC and ACSC. Through engagement, connection, and research into the application of immersive experiential learning, LII developed a learning community in this

new area incorporated by 29 different programs for enlisted, officer, and civilian education at ACSC, the Eaker Center, Air Education and Training Command (AETC), Air Combat Command, the Jeanne M. Holm Center, and the Thomas N. Barnes Center for Enlisted Education, among others.

How Experience and Mindset Are Key in Developing Faculty

A second way that LII builds and expands teaching and learning communities of practice is through faculty development, which centers around teaching and learning methods as well as course and lesson plan development that emphasize interactive and experiential leadership education experiences. LII runs faculty development annually for LDC, but it also conducts outreach faculty development for other schools and teams around AU, utilizing the models and frameworks developed through teaching in LDC. As LDC's reputation has grown, programs outside of LII and LDC have increasingly turned to LII's unique approach to faculty development, comprised of three primary frameworks that guide what the training looks like and how that training creates a larger community of teaching and learning methods across AU.

The SEE Model

The student experience ecosystem (SEE) model promotes a teaching mindset that gets away from traditional authority roles of instructor and learner in the classroom and fosters an environment of cocreation, where instructor and student are both learners and teachers with shared authority.²¹ Students take ownership of their own leadership learning by being pressed to answer their own questions or the questions of their peers instead of relying on the instructor to answer for them. The instructor is more of a facilitator who orchestrates a learning environment where students are at the center of learning and the facilitator lies more on the periphery working to create moments of introspection, learning, and

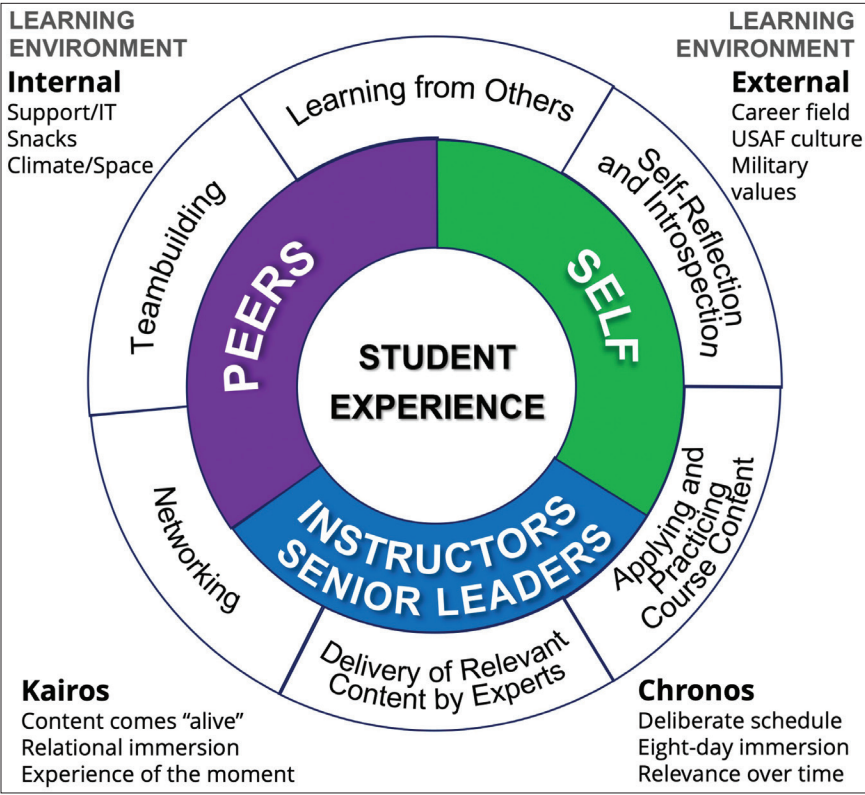
²¹ John M. Hinck and Steven B. Davis, "Re-Operationalizing and Measuring Impact of a Leader Development Course," *International Journal of Teaching and Learning in Higher Education* 32, no. 3 (2020): 427–40; and John M. Hinck and Steven B. Davis, *Going Virtual: Evolution of the Student Experience Ecosystem (SEE) Model to the SEE Model 2.0 and the Virtual (vSEE) Model*, Eaker Papers no. 4 (Maxwell Air Force Base, AL: Air University Press, 2022).

growth. This mindset is particularly suitable for leadership education like LDC and the leader education that takes place in military contexts because such training is centered more on moving the needle on individual development as opposed to teaching specific leadership models and cognitive outcomes that students are supposed to have. Rather, it focuses on introspection and self-growth, particularly in connecting to the affective domain to link the head and heart spaces for more empathetic leadership. The SEE model derives from two learning theories posited by Albert Bandura, which are social learning theory and social cognitive theory, also known as Social Cognitive Learning Theory (SCLT). Barkha Devi, Bidita Khandelwal, and Mridula Das associate Bandura's social cognitive learning theory to the process where students learn via observations from others, instructors, and the education environment created for the student.²² Through student observation with instructors, peers, and the environment, students gain knowledge with the goal of determining self-efficacy in their own leadership behaviors. The purpose of the SEE model is to create a student experience through an ecosystem of modeled behaviors versus instructing students on how to apply leadership behaviors. The SEE model aligns with Bandura's principal element of triadic reciprocity.²³ The three principles of triadic reciprocity according to Razieh Tadayon Nabavi and Devi, Khandelwal, and Das are: (1) behavior, (2) personal factors, and (3) environment factors. The triadic elements of SCLT are also a continuous reciprocal process of interaction among these three elements. The SEE model takes advantage of Bandura's concept of SCLT and the triadic reciprocity of human observation with a goal of self-efficacy of one's own behaviors. The SEE model creates observational moments through instructor modeling, peer modeling, and setting of the learning environment so that students can reflect on how others apply leadership behaviors (practice in humanity), reflect on those observational behaviors of others (instructors and peers), observe the type of environment that

²² Barkha Devi, Bidita Khandelwal, and Mridula Das, "Application of Bandura's Social Cognitive Theory in the Technology Enhanced, Blended Learning Environment," *International Journal of Applied Research* 3, no. 1 (2017): 721–24.

²³ Razieh Tadayon Nabavi, "Bandura's Social Learning Theory and Social Cognitive Learning Theory," *Theory of Developmental Psychology* 1, no. 1 (2012): 1–24.

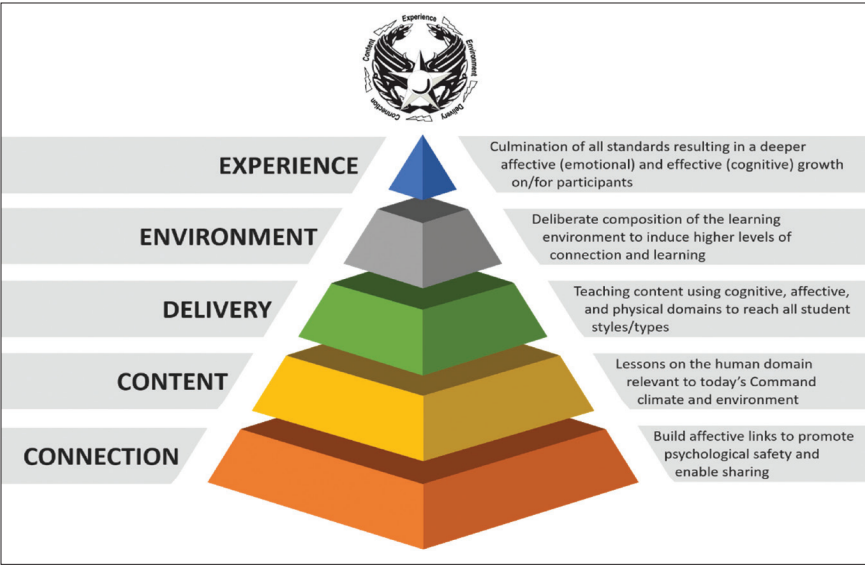
Figure 4. The student experience ecosystem (SEE) model 2.0



Source: John M. Hinck and Steven B. Davis, "Re-Operationalizing and Measuring 'Impact' of a Leader Development Course," *International Journal of Teaching and Learning in Higher Education* 32, no. 3 (2020): 427–40, adapted by MCUP.

embraces a practice in humanity, and then determine (self-reflection) when and how a student can model (change their behavior) the practice in humanity. In sum, the SEE model 2.0 (figure 4) demonstrates how students viewed LDC within the larger Air Force as an ecosystem of trust, growth, and development, with the student experience laying at the center, exemplifying Bandura's concept of SCLT and the triadic reciprocal process of learning through observation of other's behaviors, personal connectivism, and the environmental settings.

Figure 5. LDC’s pinnacle of standards



Source: John M. Hinck et al., “Understanding ‘Leaderogy’: Unique Teaching and Learning Methods in a Leader Development Course,” *Journal of Military Learning* (October 2023): 22–45, adapted by MCUP.

The Pinnacle of Standards

The second framework, the pinnacle of standards, was developed by LII as a complimentary extension of the SEE model that looked at the structure of course design as a whole. In 2020, LDC and LII were tasked with taking LDC-esque training to several Air Force major commands in abbreviated one- and two-day blocks. However, although trained LDC instructors went out to do the training, the resulting feedback was significantly lower than the norm for the LDC eight-day course. When analyzing the reason for this, LII faculty found that while instructors were delivering the same content, they were not setting the correct environmental conditions or fostering enough connection, leading to an insufficient level of student experience. From this commitment to improving learning, LII faculty established the LDC pinnacle of standards (figure 5),

which identified the student experience as resting on a foundation of fostering connection with and among participants, having relevant content, an engaging delivery of content, and orchestrating the learning environment to be conducive to experiential learning.²⁴ From this, LDC instructors were able to better tailor activities to audiences in or about to be in command; foster connection through more deliberate activities allowing for the cross-pollination of ideas and experiences in a large group settings of 30–100 people, as opposed to the typical LDC seminar of 14; and set a learning environment by rearranging tables and chairs to be more conducive to small group conversation as opposed to lecture or being talked at. In sum, the pinnacle of standards has become a central component of LII outreach and faculty development that offers other organizations a blueprint to apply in their own classrooms and learning spaces, resulting in the creation of a wider teaching and learning community at AU.

Faculty Developmental Stages

The third key area of LII's faculty development outreach is on faculty development itself, specifically when and how new instructors grow and develop and what resources and support they need during the different stages. Because LDC military instructors initially only had a 12-month assignment and taught upward of 12 courses per year, LII faculty ran a study to understand how their military instructors grew and developed and compared that to the existing literature on lifecycles of teachers, which focused on traditional teaching durations of one course per year or per semester.²⁵ The results of that study are depicted in table 3. The data showed that instructors went through the four stages of the lifecycle of teachers based not on years, which is what the previous literature talks about, but based on number of courses taught. This enables LII faculty to deliberately develop instructors to explain to them that in the survival stage of the first few courses they will feel like they are just treading water and that they should be given clear teaching plans and be allowed

²⁴ Hinck et al., "Understanding 'Leaderology'."

²⁵ Steven B. Davis and John M. Hinck, "Assessing Instructor Developmental Stages in an Accelerated Schedule: A Guided Reflection," *Journal of Faculty Development* 38, no. 3 (2024): 22–33.

Table 3. Faculty development stages

STAGE	LDC course offerings
Survival stage	1. 20–29 August
Year 1	2. 1–10 October
Year 2	3. 23–31 October
Consolidation stage	4. 12–21 November
Year 2	5. 3–12 December
Year 3	6. 7–16 January
Renewal stage	7. 28 January–6 February
Year 3	8. 18–27 February
Year 4	9. 3–12 March
Year 5	10. 24 March–2 April
Maturation stage	11. 7–16 April
Year 5	12. 28 April–7 May
Year 6	
Year 7	
Year 8+	

Source: courtesy of the athors, adapted by MCUP.

for repetition and consistency of teaching teams. As they get more comfortable, and through the consolidation stage after three to four courses, they can focus on shoring up their weaknesses and learning different teaching approaches, and by the renewal stage, typically after five or six classes, they can really start to intermix and be challenged by teaching with other instructors with strongly different approaches and personalities to expose them to new possibilities. Mixing up teaching teams allows for the cross-pollination of ideas and unique activities that they inevitably begin creating as they mature into the fourth stage, where they come into their own as an instructor and marry that identity with who they were as a leader coming into LDC. That research also showed that parallel to individual growth was the growth of the team in terms of trust, which instructors reported played a big role in their own development in learning to overcome imposter phenomenon and gain confidence and a teaching approach that was a reflection and incorporation of who they were. When the COVID-19 outbreak forced the team to convert LDC into an online course, the team reported that it was the thick trust they had built

in the previous months of teamwork that helped them overcome the fear of the unknown and successfully create a virtual LDC offering that was up to the standards that would later become the pinnacle of standards. Overall, this angle of faculty development training is particularly useful for outreach in other organizations who teach courses of different durations and frequencies.

Through these three frameworks of faculty development, LDC earned a consistent reputation as a standard of teaching and learning excellence in the leader development domain at AU. As its military instructors and course directors leave to take leadership positions in other organizations around AU and the Air Force, LII and LDC are often asked to help export LDC teaching and learning methods to those organizations. Through faculty development at these other organizations, LII helps promote and expand a teaching and learning community of practice where they are not just exporting models but engaging in conversations and learning about how they need to adapt to other organizations. In this process, LII faculty learn from these experiences and may modify the framework or how it is presented and take back other lessons learned that end up changing how LII and LDC operate. The result is a growing sense of shared practices and approaches that help propel AU away from stagnation and forward toward new possibilities.

Combined Use of Experience, Mindset, Moment, and Connections

A further example of LII's development of a teaching and learning communities of practice includes specific curriculum designed as part of LII's QEP mission. AU's QEP led LII faculty to form additional collaborative connections with faculty across AU in the development of lessons on ethical decision making and empathy. While the seeds of a learning community were already largely present within LII—a diversity of expertise, shared objectives of advancing student learning, emphasis on learning how to learn, and mechanisms for sharing lessons learned—the QEP provided greater institutional impetus and focus to such efforts, shifting LII's faculty mindset toward the purpose of AU's QEP and allowing for greater collaboration internally as well as connections made externally

through the funding of faculty development.²⁶As a result, the QEP's call to action was answered by LII's faculty who had diverse educational backgrounds—including PhDs in communication, leadership, history, and philosophy—combined with individuals with experience as PME commandants and graduated squadron commanders.

Ethical Decision Making

Through a diversity of educational and experiential backgrounds, and a commitment to rethink what effective learning on ethics looks and feels like, LII created a vibrant curriculum on ethics that significantly departed from previous and standard approaches to teaching ethics within PME, specifically those occurring outside of AU's year-long graduate programs for officers. Whereas LII's learning community contributed toward this new curriculum on ethics, its success came from more than the mere existence of a learning community to include new elements of moment, mindset, and experience.

The lynchpin of LII's ethics curriculum emerged from a key moment during initial discussions by LII faculty and instructors at LDC. In discussing what a new ethics curriculum could be, instructors identified the primary obstacle as being students' negative perceptions of ethics. This negative perspective represented a student mindset of ethics as merely a message from the top to simply "stop doing bad things" or a briefing intended to just CYA (cover your ass). Such a view was understandable in that, at best, typical PME lessons on ethics reflected instructor-oriented teaching that ostensibly encouraged students to debate various ethical dilemmas with such examples having prescriptive solutions provided or expected. At worse, such lessons were delivered half-heartedly without discussion and with content focusing on covering the rules or regulations of appropriate and inappropriate behavior; such lessons lacked authenticity by occurring only when an ethical lapse occurred. Consequently, in both content and delivery, the typical approach to teaching ethics often failed to engage students and placed them in a position of recipient rather than learner.

²⁶ Bielaczyc and Collins, "Learning Communities in Classrooms."

To foster a genuine learning community between instructor and student, LII's new ethics curriculum focused on dispelling the negative student mindset of ethics as “stop doing bad things” by moving away from a view of “ethics as compliance” to that of “ethics as aspiration.” In doing so, lessons begin with attention getters acknowledging students' typical negative mindset toward ethics and asking them what problems they see with narrowly framing ethics as solely dictated by regulations backed by punishment to deter inappropriate behavior. This, combined with examples of how such a perspective fails when regulations are unclear, yet to be created, or contradictory, in addition to challenges when oversight is limited, all serve as constructive rather than negative approaches to why ethics is important. Taken together, rather than placing agency in rules/punishment, the curriculum developed by LII anchored agency within the actions taken and decisions made into the hands of airmen.

In support of this shift in mindset, the curriculum intentionally encourages real debate and discussion over what values drive students' decision making. Instructors begin by noting the difficulty of ethics and the gray areas that students will inevitably face in a changing world, including contradictory and competing orders and mission requirements with no “answers” or solutions offered. Examples of dilemmas are purposefully crafted so that any one course of action could be justified under a specific context or set of constraints with emphasis placed on competing tensions or situational needs. Through personal reflection, discussion, and debate, students learn how others' values and justifications for their decisions depart from their own, placing the students rather than the instructor at the center of the learning environment. Students are therefore encouraged to not only offer their own examples of dilemmas but also share their personal moral values, the sources of those values, and comparison of those values with those of the Air Force, both explicitly and implicitly stated, and airmen's experiences with them.

From this shift in mindset and orchestration of a learning experience granting agency to the students' moral reasoning, LII saw success in the perceptual changes by students. Informal student feedback included statements like, “not your typical ethics lesson”; “thought provoking” and “[this] challenged me to develop my sense of ethics,” while assessment

data from AU's QEP's student surveys showed statistically significantly higher scores in students' confidence to engage in ethical thinking. Due to the success of the curriculum and because of the collaborative nature of LII's learning community and connections made with other programs across AU, LII was able to adapt and deliver this ethics curriculum to variety of schools and programs, including the Chief Leader Academy's Strategic Leader Course, Project Mercury, LDC, LeaderShop (short workshops open to Air Force audiences writ large), the U.S. Air Force Academy's National Character and Leadership Symposium, and the Barnes Center's training of PME commandants.

Empathy

To help facilitate AU's goal of fostering the concept of empathy in its airmen and guardians as part of the QEP, LII collaboratively developed and facilitated new empathy workshops starting in the spring of 2021. These workshops curated intentional learning communities where a focus on "mindset" and "connection" played a pivotal role in the learning process. Ranging from 60 minutes to 8 hours across two days, works were delivered in both online and in-person formats and were primarily developed for attendees from the Civilian Professional Development School and the Chief Leader Academy; later iterations reached AU faculty through the Teaching and Learning Center, as well as a wide range of Department of Defense employees through the Leadership and Innovation Institute's LeaderShop programming.

The overarching approach to content was to advance an understanding of empathy, and how to attune it, as well as helping individuals make the connection between empathy in their personal lives, roles as leaders, and the profession of arms. Each workshop was tailored to its specific audience, while the concepts of mindset and connection served as essential ingredients contributing to the overall participant experience. In doing so, each empathy workshop intentionally focused on participant *mindset* in three distinct phases: at the beginning, during, and after. This focus led to strong connections between participants in the overall learning community, building of trust, and meaningful reflection among groups of participants who were usually meeting each other for the first time.

In particular, at the start of each workshop, participants were invited to reflect on and share their current mindset, understanding, and experiences as they related to the concept of empathy. To get at this mindset, the questions “How do you define empathy?” and “What is your relationship to it?” were posed. The motivation here was not about gaining or sharing accurate information, *per se*—in fact, much of what was shared was often in opposition to the academic literature on empathy. However, this initial brainstorming and sharing session fostered connections between workshop participants in their new learning environment prior to getting into content. As personal examples and stories were shared and affirmed, they formed the building blocks of commonalities and divergences between participants, reflecting the notion of “connection before content” as key to mindset.

Next, LII faculty leveraged this supportive environment through content delivery. The “five considerations of empathy,” which include: 1) empathy is emotional perspective taking, 2) empathy involves both emotion and cognition, 3) empathy requires individuals to recognize the feelings they “catch” are not their own, 4) empathy can be learned, and 5) empathy is best understood as a continuum instead of a categorical concept, were the basis of most workshops. These assumptions were often contrary to students’ prior experiences and tended to be substantially different than the ideas about empathy shared in the opening phase. The results, evidenced by regular feedback from participants after the workshop—both formally through QEP surveys, but also informal follow up in Zoom chat, via email to instructors, or face-to-face—confirmed that participants felt as if they could share and grow in the space curated.

Finally, each workshop concluded (the “after phase”) with a deliberate refocus on participant mindset and growth. Attendees were asked to share their progress and understanding as it related to their position at the start of the workshop. The question “On a scale of 1–5, how would you describe your growth as it relates to empathy?” was posed, followed by an open discussion. Feelings of growth, change, and even shock were often themes of these outbriefs, with participants frequently commenting on the inclusive nature of the workshop and the connections they made

during their session, Informally, many exchanged contact information with the group and/or individuals they met.

By actively engaging participant mindset before, during, and after content, the workshops led to supportive learning environments through which participants were able to develop their empathy knowledge and skills. As evident by AU's QEP assessment data, students' self-perception and knowledge of empathy significantly increased in all nine question categories asked. Taken together, the empathy workshops developed and facilitated as part of QEP efforts beginning in 2021 contributed to intentional learning communities across AU. With thoughtful focus on mindset and fostering connections prior to content, participants were able to not only grow their understanding of empathy, but also their relationship to learning.²⁷

Postvention Training

A third content area where LII formed a teaching and learning community is in postvention. Postvention, defined as activities developed by, with, or for suicide survivors, to facilitate recovery after suicide, and to prevent adverse outcomes including suicidal behavior, focuses on responding to a suicide death by helping people grieve, minimizing blame, and ensuring safe messaging.²⁸ The goals of postvention are to promote healthy healing, reduce risk of contagion, and link others at risk to resources.²⁹ While existing postvention programs focus on specific settings, such as schools, workplaces, and faith communities, it is particularly important to implement postvention efforts in the military due to suicide being one of the

²⁷ Antinluoma, Ilomaki, and Toom, "Practices of Professional Learning Communities," 1-14; Bielaczyc and Collins, "Learning Communities in Classrooms"; and DuFour, "What Is a Professional Learning Community," 9.

²⁸ *Preventing Suicide in the U.S. Military: Recommendations from the Suicide Prevention and Response Independent Review Committee* (Washington, DC: Suicide Prevention and Response Independent Review Committee, Department of Defense, 2022).

²⁹ *Annual Suicide Report: Calendar Year 2020* (Washington, DC: Defense Suicide Prevention Office, Under Secretary of Defense for Personnel and Readiness, Department of Defense, 2021).

leading causes of death in the military.³⁰ Here, LII created and teaches curriculum utilizing connection and experience to deliberately develop the mental health of airmen, enabling them to better lead during the loss of life in a post-suicide crisis response incident.

When communities lose a member to suicide, the experience has a socioecological impact across multiple domains.³¹ Within a military community, suicide impacts the morale and mission of its servicemembers, making it crucial that leaders can effectively apply evidence-based concepts to return an organization to a healthier operational presuicide status.³² In support of this, LII's postvention training and education enables leaders to develop the mindset and skills to better communicate care while accomplishing their purpose and mission. In doing so, LII leveraged its resident postvention expert, Dr. Mary Bartlett, to collaboratively engage both faculty and students across the Department of the Air Force, most notably through the new postvention curriculum in LDC.

Responding to consistent student feedback during the course of LDC's initial years of executions, LII faculty introduced and continuously revised segments on crisis and postvention through the incorporation of storytelling, scenarios, and large group discussions. In addition to providing an explanation of postvention and education about its application in reducing the likelihood of future suicide, LDC's postvention curriculum takes a student-centered approach emphasizing shared learning and emotional connection modeled first by LDC instructors and graduated squadron commanders' own stories of leading through and being impacted by postvention to generate a culture of trust to emerge, contributing toward students' sharing of their own stories of postvention.

In addition to facilitating deep, candid, robust, and often challeng-

³⁰ Marie-Louise Sharp et al., "Stigma as a Barrier to Seeking Health Care among Military Personnel with Mental Health Problems," *Epidemiologic Reviews* 37, no. 1 (2015), <https://doi.org/10.1093/epirev/mxu012>.

³¹ J. R. Jordan, "Grief After Suicide: The Evolution of Suicide Postvention," in *Death, Dying, and Bereavement: Contemporary Perspectives, Institutions, and Practices*, ed., J. M. Stillion and T. Attig (New York: Springer, 2015), 349–62.

³² Tiffany M. Greene-Shortridge, Thomas W. Britt, and Carl Andrew Castro, "The Stigma of Mental Health Problems in the Military," *Military Medicine* 172, no. 2 (2017): 157–61, <https://doi.org/10.7205/MILMED.172.2.157>.

ing discussions about the postvention, LII's LDC curriculum offers an experiential approach by incorporating mixed reality scenarios to engage students in all domains of learning (cognitive, affective, and behavioral). Here, students engage in conversations with an avatar also trained in postvention to create an authentic experience to better prepare airmen for the challenges of command. This incorporation of authentic role-playing connects the students emotionally and physically in a realistic simulation followed by students debriefing their experiences through a group conversation to explore the response processes, identify alternate and better response options, discuss observations and shared experiences, and ultimately connect them to their own realities, biases, and support network should they lead through a suicide attempt and/or completion. In this context, participants glean information on how to successfully achieve the core goals of effective postvention to foster healthy grieving, reduce the likelihood of contagion, and identify efficacious resources to help lead through a complex, emotionally draining, and potentially growth-enhancing—though tragic—situation.

In sum, LII's efforts led to LDC becoming one of the first Air and Space Forces PME organizations to formally include postvention into its curriculum with specific lesson objectives created, lesson plans developed, and faculty and instructor training formalized. LII's teaching and learning community enabled such efforts, resulting in LDC being among the first courses to also quantify outcomes for this framework, while taking the content to a more innovative and higher level of learning. Taken together, this combination of content, experience, and reflection allows for students to prepare and develop a mindset and experiential understanding to better lead during postvention; one demonstrated to produce more ethical, empathetic, and prepared leaders across the Air and Space Forces enterprise.³³ As human beings continue to wrestle with the topic of suicide, the challenge of how best to educate, train, and develop military leaders and others to appropriately respond to the actualities of suicide before, during, and after such events is vitally important. By inviting

³³ Hinck and Davis, "Re-Operationalizing and Measuring Impact," 427–40; and Hinck and Davis, "Going Virtual."

and providing leaders-in-training a safe space to discuss mental health and post-suicide experiences often not discussed, stigma can be reduced, and lives can be saved.

Implications, Applications, and Limitations

As the Air Force optimizes its education and training for great power competition, examining alternative teaching and learning processes is necessary. This case study explored one such model, that of a teaching and learning community of practice. As a community dedicated to deep learning, not simply teaching or disseminating “turnkey” content, LII has impacted leadership learning and teaching across AU in novel and impactful ways. From this approach, LII’s efforts have already enacted some elements of AU’s strategic plan, including key focus areas of “Enhance Force Development Delivery” and “Modernize the Environment,” in addition to AETC’s operational approaches of “Pivot to Tech Training” and “Incorporate Digital-Age Technology.”³⁴ LII’s success in these areas suggest other PME organizations, large and small, may find the learning community approach to PME more impactful in delivering quickly adaptive and impactful education and training by understanding that an organization’s TLCoP: 1) comprises multiple concepts and multiple levels; 2) should focus on key concepts that bridge experience, mindset, moments, and connection; 3) view and approach teaching and learning that are inclusive of all participants (learners and teachers) in a holistic way; and 4) consider a holistic participant-centered community approach that is anchored in collaboration and key learning outcomes. It may be worthwhile to show how LII overcame some of the limitations (obstacles and challenges) to achieve the great successes promoted throughout this chapter. While a lot has been accomplished, it has not been without overcoming challenges. Any time work of this nature is executed, particularly within the human domain and mental health realms, obstacles exist that must be pushed through. Many people are unfamiliar with the multitude

³⁴ *Strategic Plan for 2014–2020* (Birmingham, AL: Air University, 2013); and *Strategic Communication Plan for Fiscal Year 2021* (Birmingham, AL: Air Education and Training Command, 2020).

of educational goals, competing needs, directives, governing policies, and politics that AU works within. As the epicenter of learning for the entire Air Force, the enterprise is not just accountable for premier academic outcomes but the enormous responsibility of producing some of the world's greatest leaders in air power. LII is one of many PME products seeking to set itself apart in making an impact.

All the people involved in these projects, all the groups and different agencies, must approve them, provide inputs, use them, and receive them, and none of these people are coordinated with each other. They each have their own aims and goals and are mutually exclusive. So, the reasonable person may believe it is near impossible to produce a single product. What these tools provide is the possibility to provide some sort of positive benefit that the products need to provide flexibility. They are created and implemented in such a way so that flexibility is preserved. What is required from leadership is the capacity to let go, provide overall guidance, and allow creative flexibility. Yet, often in academics, standardization is sought that complicates matters, especially in the military where standardization is highly valued.

The leadership development process takes decades to achieve, and one obstacle can be the perception that these areas can be developed in a short time. You can consider ideas, account for the effect of new technology, reflect on your own mistakes, and explore your own biases, and that sort of self-realization is an evolution. The skills proposed in this context were often met with institutional inertia, which required faculty to not just think creatively but compromise and require that their new ideas be subject to some sort of reasonable oversight rather than merely embracing a new idea because it is new. To achieve these successes, implementation of these ideas, models, and skills required dedicated perseverance and a certain degree of patience with bureaucratic drag.

Conclusion

This chapter examined LII as a case study of a PME organization approaching teaching and learning through a TLCoP. Results of the cumulative coding process found LII's TLCoP comprising of three levels (primary, secondary, and tertiary), with the primary level including six

constructs based in experience, mindset, moments, connection, collaboration, and participant as both learner and teacher (shared authority). Importantly, whereas analysis of LII's TLCoP demonstrated typical variables associated with TLCoP, four additional constructs of experience, mindset, moments, and connection emerged. These new constructs help further explain the means by which effective TLCoP can cocreate effective learning processes, as shown by LII's efforts involving mixed reality/innovation and faculty development, in addition to specific curricula on ethics, empathy, and postvention training. In sum, this chapter provides timely, relevant, and impactful higher educational strategies for teaching and learning by unpacking the creative and innovative immersive technology and experiences being applied across AU, which, when taken together, result in highly affective and experiential approaches to the development of leadership and educational development for faculty, staff, instructors, and learners across the Air Force enterprise.



CHAPTER 2 ●

What Do Creative Problem-Solvers Need?

Lieutenant Colonel Timothy A. Sparks (USMC), EdD

In military contexts, a tension exists between the need for rapid, unquestioning obedience to orders, especially early in one's career, and the need for senior leaders to solve complex problems creatively. For officers in the U.S. Marine Corps, a key milestone in their careers is the Marine Corps' Command and Staff College (CSC), an intermediate-level professional military education (PME) master's degree program. Using a convergent parallel mixed-methods design, this study examined factors related to creative problem solving and their application to the CSC curriculum. To facilitate this examination, a survey was developed and tested to examine selected factors in the hope that it might prove useful for this context and other related military contexts.

Background

Military servicemembers face an unpredictable and uncertain world that requires them to think creatively to solve complex problems.¹ In a sense, this is nothing new. The Marine Corps' seminal publication, *Warfighting*, makes it clear that the purpose of PME is "to develop creative, thinking

¹ Gen David H. Berger, *Commandant's Planning Guidance: 38th Commandant of the Marine Corps* (Washington, DC: Headquarters Marine Corps, 2019), 16; and *Developing Today's Joint Officers for Tomorrow's Ways of War: The Joint Chiefs of Staff Vision and Guidance for Professional Military Education & Talent Management* (Washington, DC: Joint Chiefs of Staff, 2020), 3.

leaders.”² At the same time, many leaders are convinced that the U.S. PME community needs to be reinvigorated.³ According to one compilation of studies on military creativity, “There is an inherent tension between encouraging creativity within the Armed Forces and maintaining military discipline.”⁴ Students at CSC find themselves in a transition to increased responsibility and a greater need for enhanced creative problem solving as they learn to plan at the operational level of war.

To meet this need, Marine Corps University (MCU) developed a quality enhancement plan to assess and improve demonstrated creative problem-solving ability in student products. Table I reports the results of these assessments for CSC. The university used Teresa M. Amabile’s consensual assessment technique to rate the creativity of student products.⁵ The consensual assessment technique employs subject matter experts to evaluate the creativity of products.⁶ A particular strength of the consensual assessment technique is that it looks at the products and not the personality or cognitive style of the creators.⁷ Importantly, the technique employs experts in the field, who focus on the creativity of the products and not necessarily on technical matters such as writing style and grammar.⁸

The use of the consensual assessment technique has precedent in the military education community. Prior to this study, at least three other studies used the consensual assessment technique to evaluate the creativity of products created by Australian Air Force officers and U.S. Army

² *Warfighting*, Marine Corps Doctrinal Publication 1 (Washington, DC: Headquarters Marine Corps, 1997), 61.

³ Berger, *Commandant’s Planning Guidance*, 16; and *Developing Today’s Joint Officers for Tomorrow’s Ways of War*, 2.

⁴ Susan F. Bryant and Andrew Harrison, *Finding Ender: Exploring the Intersections of Creativity, Innovation, and Talent Management in the U.S. Armed Forces* (Washington, DC: National Defense University Press, 2019), 1.

⁵ Teresa M. Amabile, “Social Psychology of Creativity: A Consensual Assessment Technique,” *Journal of Personality and Social Psychology* 43, no. 5 (1982): 997, <https://doi.org/10.1037/0022-3514.43.5.997>.

⁶ Amabile, “Social Psychology of Creativity,” 1006.

⁷ Amabile, “Social Psychology of Creativity,” 1001.

⁸ Amabile, “Social Psychology of Creativity,” 1010.

officers in a professional context.⁹ Another study, conducted in a U.S. civilian graduate school context with a multicultural cohort of students, also used the consensual assessment technique to evaluate final papers after exposure to multicultural dialogue in a semester-long course.¹⁰

To ensure the reliability of the quality enhancement plan's assessment results, the university augmented Amabile's technique with the American Association of Colleges and Universities' (AAC&U) creative-thinking value rubric, which identifies the following six components of creative thinking: acquiring strategies and skills, embracing contradictions, connecting and synthesizing, innovative thinking, solving problems, and taking risks.¹¹ For each of these components, the rubric provides criteria to assess the students' work on the following scale: not demonstrated (0), imitative (1), adaptive (2), creative (3), and transformative (4).¹² After each rater provided a score for each component, the scores were then averaged and the result was rounded to the nearest whole number. For student products rated by more than one rater, the two rater's scores were averaged. MCU's goal was that 80 percent of students would score as creative or transformative (2.5 or above on the scale). In the MCU context, the products assessed ranged from student essays submitted for existing class assignments to student presentations at staff rides.¹³

The assessments for academic years ending in 2015 through 2020 are

⁹ LtCol Richard A. McConnell et al., "Improving Creative Thinking through Narrative Practice," in *Developments in Business Simulation and Experiential Learning: Proceedings of the Annual ABSEL Conference*, vol. 50 (Fort Leavenworth, KS: Army University Press, 2023), 341–94; Bree L. Sandwith, David H. Cropley, and Lisa J. Chantler, "The Influence of Cognitive Structure and Task Structure on Creativity in a Military Context," *International Journal of Creativity and Problem Solving* 27, no. 2 (2017): 95–112; and Andrea S. Vincent, Brian P. Decker, and Michael D. Mumford, "Divergent Thinking, Intelligence, and Expertise: A Test of Alternative Models," *Creativity Research Journal* 14, no. 2 (2002): 163–78, https://doi.org/10.1207/S15326934CRJ1402_4.

¹⁰ Jennifer L. Matic, "Multicultural Student-to-Student Dialogue and the Generation of Creative Ideas," *College Teaching* 67, no. 2 (2019): 86–93, <https://doi.org/10.1080/87567555.2018.1558168>.

¹¹ *Creative Thinking VALUE Rubric* (Washington, DC: American Association of Colleges and Universities, 2009).

¹² *Creative Thinking VALUE Rubric*, 2.

¹³ *Fifth Year Interim Report*, pt. 5, *QEP Impact Report*. (Quantico, VA: Marine Corps University, 2021), 5.

Table 1. Percent of student products rated as creative or transformative

AY15		AY16		AY17		AY18	
Fall	Spring	Fall	Spring	Fall	Spring	Fall	Spring
7%	13%	69%	81%	1%	13%	3%	4%

Source: courtesy of the author.

reported in MCU’s impact report of the quality enhancement plan. For academic year 2016, the impact report indicates that the university used a different assessment tool and that the results for that year should not be compared to the other year’ results. After the impact report was complete, and though no longer required, the university continued to assess student creative problem solving using the procedures established for the quality enhancement plan. Overall, in every year except academic year 2016, CSC did not meet its goal of 80 percent of students producing work that was creative or transformative. This study aims to discover why the university’s goals were not met.¹⁴

Following Amabile’s model, MCU focused on the usefulness of products, not necessarily the creativity of the individuals who made them.¹⁵ Therefore, MCU refined the consideration of creativity, which may manifest in multiple ways, to a consideration of how creative thought is applied in practical solutions to problems of import to the national security of the United States and its allies.¹⁶ Creativity is more than just novel ideas—these ideas must also be useful and be able to be put into practice.¹⁷ Creative problem-solving takes creativity one step further by requiring that students be able to provide a plan for implementation to turn an abstract idea into a functional process, including enough detail to be able to be put into practice.

¹⁴ *Fifth Year Interim Report*, 5.
¹⁵ *Strengthening Leadership through Enhanced Creative Problem Solving: A Quality Enhancement Plan for Marine Corps University, 2015–2020* (Quantico, VA: Marine Corps University, 2021), 30; and Amabile, “Social Psychology of Creativity,” 1010.
¹⁶ *Strengthening Leadership through Enhanced Creative Problem Solving*, 20.
¹⁷ *Strengthening Leadership through Enhanced Creative Problem Solving*, 27.

AY19		AY20		AY21		AY22	
Fall	Spring	Fall	Spring	Fall	Spring	Fall	Spring
4%	8%	39%	49%	40%	45%	52%	n/a

Table 2. Components of creativity

Components	Specific constructs examined
Creative thinking skills	Divergent thinking Metacognition Authoritarianism
Skills and knowledge related to the task	Subject matter knowledge Research literacy
Motivation/opportunity	Intrinsic motivation Time

Source: courtesy of the author.

Conceptual Framework

The conceptual framework for this study draws inspiration from Amabile’s componential model of creativity in individuals and innovation in organizations. According to Amabile’s model, there are three main components to creativity in individuals, with three related components for innovation in organizations. Amabile’s three main components for individuals are skills and knowledge related to the task, motivation, and creative-thinking skills, while components at the organizational level relate to innovation management and providing proper motivation and opportunity.¹⁸

¹⁸ Teresa M. Amabile and Michael G. Pratt, “The Dynamic Componential Model of Creativity and Innovation in Organizations: Making Progress, Making Meaning,” *Research in Organizational Behavior* 36 (2016): 161, <https://doi.org/10.1016/j.riob.2016.10.001>.

Within the component of creative-thinking skills, this study investigated how divergent thinking and metacognition contribute to creative problem-solving in a military context. *Divergent thinking* refers the ability to approach problems from alternative perspectives and develop alternative solutions.¹⁹ This ability is a vital skill for creative thinking. However, some scholars argue that military culture stifles divergent thinking. For example, Susan F. Bryant and Andrew Harrison argue that obedience is more important than divergent thinking when it comes to military talent management.²⁰ Moreover, two studies focused on the U.S. Military Academy at West Point and the U.S. Naval Academy found that divergent thinkers are less likely to persist in these institutions.²¹ Additionally, other studies found little improvement when officers reached senior-level leadership.²² Conversely, other studies have found that the flexibility required in military life translates to improved divergent thinking skills.²³

Metacognition refers to the ability to evaluate personal knowledge and the capacity to strategize ways of improving it.²⁴ Research suggests that metacognition is an effective strategy for building critical and creative

¹⁹ Vincent, Decker, and Mumford, "Divergent Thinking, Intelligence, and Expertise," 163.

²⁰ Bryant and Harrison, *Finding Ender*, 7.

²¹ Tom Mitchell and Alice M. Cahill, "Cognitive Style and Plebe Turnover at the U.S. Naval Academy," *Perceptual and Motor Skills* 101, no. 1 (2005): 55–62, <https://doi.org/10.2466/pms.101.1.55-62>; and Angela Duckworth, *Grit: The Power of Passion and Perseverance* (New York: Scribner, 2016), 3–8.

²² Kimberly Jackson et al., *Raising the Flag: Implications of U.S. Military Approaches to General and Flag Officer Development* (Santa Monica, CA: Rand, 2020), 181, <https://doi.org/10.7249/RR4347>; and Leonard Wong and Stephen J. Gerrass, *Changing Minds in the Army: Why It Is So Difficult and What to Do about It* (Carlisle, PA: U.S. Army War College Press, 2013), ix.

²³ Stephen J. Zaccaro et al., "The Influence of Higher Order Cognitive Capacities on Leader Organizational Continuance and Retention: The Mediating Role of Developmental Experiences," *Leadership Quarterly* 26, no. 3 (2015): 352, <https://doi.org/10.1016/j.leaqua.2015.03.007>; and Bree L. Sandwith, David H. Cropley, and Lisa J. Chantler, "The Influence of Cognitive Structure and Task Structure on Creativity in a Military Context," *International Journal of Creativity and Problem Solving* 27, no. 2 (2017): 108.

²⁴ Angelle A. Khachadoorian, Susan L. Steen, and Lauren B. Mackenzie, "Metacognition and the Military Student: Pedagogical Considerations for Teaching Senior Officers in Professional Military Education," *Journal of Military Learning* (April 2020): 6.

thinking skills.²⁵ John H. Flavell introduced the idea of metacognition with a focus on how children learn, but also theorized that the same general processes are involved in adult learning.²⁶ In general, metacognition refers to two basic capacities: the capacity to evaluate current knowledge and the capacity to strategize ways of improving it.²⁷ Some research indicates that students might benefit from metacognitive strategies to encourage divergent thinking and reflection, leading to increased creativity, which made metacognition a special area of interest.²⁸

A potential barrier to creative problem solving, especially to the influence of divergent thinking, is *military authoritarianism*, which is the preference for solutions deriving from existing authority and the deference to superior ranking officers due to a reticence to question their decisions and opinions.²⁹ Balancing respect for military authority and critical thinking can be difficult, both for military students and for military faculty, and power disparities between junior personnel and senior officers might cause junior personnel to avoid communicating critical or creative thoughts to their commanders.³⁰ Moreover, another study found strong biases toward the status quo and hierarchical decision-making in the

²⁵ Roy Y. J. Chua, Michael W. Morris, and Shira Mor, "Collaborating across Cultures: Cultural Metacognition and Affect-based Trust in Creative Collaboration," *Organizational Behavior and Human Decision Processes* 118, no. 2 (2012): 128, <https://dx.doi.org/10.1016/j.obhdp.2012.03.009>; and Kelly Y. L. Ku and Irene T. Ho, "Metacognitive Strategies that Enhance Critical Thinking," *Metacognition and Learning* 5, no. 3 (2010): 264, <https://doi.org/10.1007/s11409-010-9060-6>.

²⁶ John H. Flavell, "Metacognition and Cognitive Monitoring: A New Area of Cognitive-developmental Inquiry," *American Psychologist* 34, no. 10 (1979): 908–10, <https://doi.org/10.1037/0003-066X.34.10.906>.

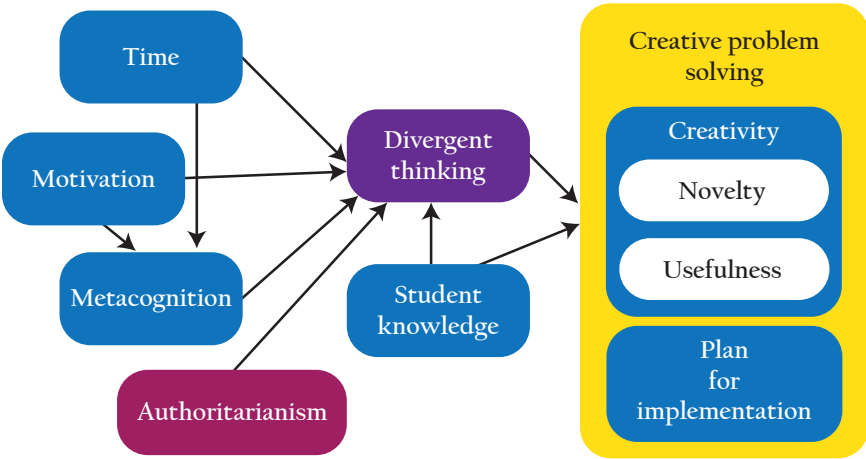
²⁷ Khachadoorian, Steen, and Mackenzie, "Metacognition and the Military Student," 4; and Flavell, "Metacognition and Cognitive Monitoring," 906.

²⁸ Chua, Morris, and Mor, "Collaborating Across Cultures," 116–31; Khachadoorian, Steen, and Mackenzie, "Metacognition and the Military Student," 3–18; and Ku and Ho, "Metacognitive Strategies," 251–67.

²⁹ Andrew Hill, "Military Innovation and Military Culture," *Parameters* 45, no. 1 (2015): 86–98, <https://doi.org/10.55540/0031-1723.2809>; Bryant and Harrison, *Finding Ender*, 6–10; and Wong and Gerras, *Changing Minds*, 20–23.

³⁰ Wong and Gerras, *Changing Minds*, 22; Danic Parenteau, "Teaching Professional Use of Critical Thinking to Officer-Cadets: Reflection on the Intellectual Training of Young Officers at Military Academies," *Journal of Military Learning* (April 2021): 50–53.

Figure 1. Conceptual framework



Source: courtesy of the author, adapted by MCUP.

U.S. Navy.³¹ Therefore, the study also considered the possibility that authoritarian attitudes might serve as a barrier to creative problem solving, hypothesizing a negative relationship between divergent thinking and authoritarianism (shown in red in figure 1).

Within the category of skills and knowledge related to the task, the study focused primarily on students' knowledge of the subject matter at CSC, including research literacy. Finally, within the category of motivation and opportunity, the study considered the motivation of the students as well as the resource of time. The hypothesized relationships between each of these constructs is diagrammed below.

Research Questions

1. To what degree do students in a PME context:
 - a. Perceive the environment to encourage divergent thinking;

³¹ Christoffer Eriksen, "On Creativity: A Case Study of Military Innovation" (master's thesis, Naval Postgraduate School, 2015), 84.

- b. Appeal to authoritarianism or perceive the environment to be authoritarian;
 - c. Perceive that they have the time necessary to think creatively;
 - d. Perceive that they have the requisite subject matter knowledge and research literacy;
 - e. Practice metacognition; and
 - f. Demonstrate motivation to think creatively?
2. To what degree do faculty in a PME context perceive that:
 - a. The environment encourages divergent thinking;
 - b. The environment is authoritarian;
 - c. Students have the necessary subject matter and research literacy;
 - d. Students have the time necessary to think creatively; and
 - e. Students are motivated to think creatively?
3. How do students perform on assessments related to creative problem solving, and how do these assessments relate to the above factors?
4. How do faculty and students in a PME context describe the environment for creative problem solving at MCU?

Population

For year one of this study, focused on secondary student data, the population was 213 mid-career military officers and interagency civilian federal employees. These students included 106 U.S. Marines, 60 other U.S. military officers, 30 international military officers, and 17 civilian federal employees. For year two of the study, the population was similar (n=209). Slightly more than one-half (n=110) of the students were Marine officers, while the others were a mixture of other U.S. Service officers (n=52), interagency civilians (n=15), or international military officers (n=32). The faculty population was 60 faculty members who lectured, led seminars, or mentored theses for students at CSC. One-half (n=30) of these faculty were civilians. The remainder of the faculty were military officers who were usually senior in rank and in experience to the students (n=27) or

interagency civilians with experience in other elements of the federal government (n=3).³²

Methodology

This project uses a convergent parallel mixed methods design with two primary methods of data collection: survey methodology and secondary data analysis.³³ These two methods of data collection yielded both quantitative and qualitative data. Data sources included an annual MCU student survey (quantitative and qualitative), new surveys of faculty and students developed specifically for this project (quantitative and qualitative), interviews to refine these new surveys (qualitative), an assessment of student creative problem-solving conducted for the quality enhancement plan (quantitative), publicly available master's theses (quantitative), and student grades (quantitative). Secondary data originated primarily from year one of the study, while the new surveys developed specifically for this project (and associated grading data) were collected from year two of the study.

This project did not reevaluate the creativity of any student artifacts but drew on the existing evaluations of students' work as reflected in grades or conducted as a part of the assessment of the quality enhancement plan. While the newly developed faculty survey was anonymous, the new student survey gave students the option of providing their name so that the researcher could compare their survey responses with quality enhancement plan (QEP) assessment scores, grades on select assignments, and first semester grade point averages (GPA).

Survey Development

The purpose of using surveys in this study was to determine to what extent factors identified in the conceptual framework are present in students and their context at MCU. Ultimately, the results of the surveys should inform curriculum innovation to foster creativity in MCU stu-

³² *Marine Corps University Factbook 2022* (Quantico, VA: Marine Corps University, 2022), 18–19.

³³ John W. Creswell and Vicki L. Plano Clark, *Designing and Conducting Mixed Methods Research*, 2d ed. (Los Angeles, CA: Sage Publications, 2011), 246.

dents and potentially across the PME community. By choosing to conduct surveys instead of focus groups or interviews, the researcher focused primarily on their utility for deductive studies, in that the factors involved derive from Amabile's work on creativity and the intent was to determine if those same factors exist at MCU.³⁴ Additionally, surveys are useful for this project based on their utility for efficiently gathering information about a large population.³⁵

One disadvantage of surveys is that they are more difficult to adapt to unique contexts like that of the U.S. military.³⁶ For this reason, the survey instrument for this study is primarily new. Though all questions are supported by the conceptual framework, some are also inspired by other surveys about fostering creativity in higher education, including graduate higher education.³⁷ Because the new surveys have been tailored to the MCU environment, if they prove to be valid and reliable, they might later be tested for use at other U.S. military institutions of higher education.

The survey questionnaires used in this study include a mix of demographic questions, multiple-choice questions, and open-ended questions.³⁸ The multiple-choice questions for both surveys used a five-point

³⁴ Chad Lochmiller and Jessica N. Lester, *An Introduction to Educational Research: Connecting Methods to Practice* (Los Angeles, CA: Sage Publications, 2015), 34; and Amabile and Pratt, "The Dynamic Componential Model of Creativity and Innovation in Organizations," 161.

³⁵ Drew Allen, "Surveys," in *Research in the College Context: Approaches and Methods*, 2d ed., ed., Frances K. Stage and Kathleen Manning (New York: Routledge, 2016), 157.

³⁶ Allen, "Surveys," 163.

³⁷ Eunice Maria Lima Soriano de Alencar and Denise de Souza Fleith, "Criatividade na Educação Superior: Fatores Inibidores," *Avaliação: Revista da Avaliação da Educação Superior (Campinas)* 15, no. 2 (2010): 201–19, <https://doi.org/10.1590/S1414-4077201000020001I>; Eunice M. L. Soriano de Alencar and Denise de Souza Fleith, "Inventory of Teaching Practices for Creativity in Higher Education," in *Theory and Practice of Creativity Measurement* (New York: Routledge, 2014), 51–64; Denise de Souza Fleith and Cristiano Mauro Assis Gomes, "Students' Assessment of Teaching Practices for Creativity in Graduate Programs," *Avaliação Psicológica: Interamerican Journal of Psychological Assessment* 18, no. 3 (2019): 306–15, <https://doi.org/10.15689/ap.2019.1803.15579.10>; and Ellie Melissa Potter, "Perceptions of Creativity among Faculty in Higher Education" (master's thesis, Grand Valley State University, 2013).

³⁸ For a list of the survey questions, see Timothy A. Sparks, "Fostering Graduate Student Creative Problem Solving in a Professional Military Education Context" (EdD diss., Johns Hopkins University, 2023), 203–6.

Likert scale for responses.³⁹ Each survey questionnaire included four to five questions related to each of the constructs identified in the first two research questions, except that metacognition was only overtly present in the student survey. Additionally, open-ended questions were included to give respondents a chance to provide additional details about their opinions.⁴⁰

Eunice M. L. Soriano de Alencar and Denise de Souza Fleith developed two surveys in the higher education context to determine what factors faculty consider to be barriers to creativity and which creativity-related pedagogical practices students report that their faculty use.⁴¹ To better fit the MCU context and the models above, this study did not adopt those surveys directly, but it drew inspiration for many of its questions from these two surveys. All questions were significantly reworded. Additionally, this project derived inspiration from Ellie Melissa Potter's open-ended survey of higher education faculty concerning barriers and aids to student creativity.⁴²

The questions related to metacognition, unlike the others, were drawn directly from Gregory Schraw and Rayne Sperling Dennison's metacognitive awareness inventory. The two factors that Schraw and Dennison reported were knowledge of cognition and regulation of cognition. For the current study's student survey, from the 52-item metacognitive awareness inventory, three questions were selected from each factor based on their theoretical relevance to creative problem solving.⁴³

Reliability and Validity of the New Surveys

Researchers responsible for each of the surveys that inspired the present study's questions conducted their own reliability and validity measure-

³⁹ Allen, "Surveys," 163.

⁴⁰ Allen, "Surveys," 163.

⁴¹ Alencar and Fleith, "Criatividade na Educação Superior," 208; and Alencar and Fleith, "Inventory of Teaching Practices for Creativity in Higher Education," 60–62.

⁴² Potter, "Perceptions of Creativity among Faculty in Higher Education," 49–50.

⁴³ Gregory Schraw and Rayne Sperling Dennison, "Assessing Metacognitive Awareness," *Contemporary Educational Psychology* 19, no. 4 (1994): 473–74, <https://doi.org/10.1006/ceps.1994.1033>.

ments.⁴⁴ Because the current versions of the faculty and student surveys are amalgamated from questions from multiple researchers and multiple surveys, validity needed to be reassessed.⁴⁵ Additionally, a confirmatory factor analysis was conducted for each of the theoretical constructs.⁴⁶

As an initial test for the validity of the questions on the faculty survey questionnaire, the researcher conducted cognitive interviews with four participants who were either current or former faculty members at MCU, of whom one was a woman and one was African American. All others were White men. At the time of the interview, two were civilians and two were military officers. Similarly, as an initial test for the validity of the questions on the student survey questionnaire, the researcher conducted cognitive interviews with five military officers who had graduated from intermediate-level PME schools. Among the officers interviewed for the student survey, three were Hispanic men, one was a White woman, and one was a White man.

Cognitive interviews used the think-aloud method to increase the validity of the questions, thereby making sure that questions were worded clearly and that participants understood them as intended.⁴⁷ After each interview, the researcher revised the applicable questionnaire or made comments on the questionnaire to address later, sometimes asking later participants to weigh in on recommendations of earlier participants.

Following completion of the cognitive interviews, the researcher made several changes to the question wording on each questionnaire to clarify the meaning of the questions. For example, one question asked about learning strategies, but military students and faculty who partici-

⁴⁴ Alencar and Fleith, "Criatividade na Educação Superior," 208–9; Alencar and Fleith, "Inventory of Teaching Practices for Creativity in Higher Education," 57–59; Fleith and Gomes, "Students' Assessment of Teaching Practices for Creativity in Graduate Programs," 306; Potter, "Perceptions of Creativity among Faculty in Higher Education," 38; and Schraw and Dennison, "Assessing Metacognitive Awareness," 460.

⁴⁵ Laura M. Desimone and Kerstin Carlson Le Floch, "Are We Asking the Right Questions?: Using Cognitive Interviews to Improve Surveys in Education Research," *Educational Evaluation and Policy Analysis* 26, no. 1 (2004): 1, <https://doi.org/10.3102/01623737026001001>.

⁴⁶ Alan Agresti and Barbara Finlay, *Statistical Methods for the Social Sciences*, 4th ed. (Upper Saddle River, NJ: Pearson Prentice Hall, 2009), 533.

⁴⁷ Desimone and Le Floch, "Are we Asking the Right Questions?," 4.

pated in the cognitive interview expressed confusion about the use of the word “strategies.” The word “strategies” implies, in the military context, a much wider scope of responsibility than was intended by the question. As a result, the word “strategies” was replaced in the final version of the survey with the word “tactics,” which implies action taken at the individual or small group level.

Confirmatory Factor Analysis

Student surveys included the following theoretical factors: divergent thinking, authoritarianism, student knowledge, time and opportunity, motivation, and metacognition. Consistent with previous survey research on metacognition, the metacognition factor was further divided into knowledge of cognition and regulation of cognition.⁴⁸ The initial results of factor analysis showed a significant chi-squared test of model fit ($\chi^2(362, N = 66) = 572.07, p < .0001$). The root mean square error of approximation was 0.094 ($p < .0001$), which was a mediocre fit. The Comparative Fit Index (CFI) (0.662) and Tucker-Lewis Index (TLI) (0.621) were weak. Finally, the Standardized Root Mean Square Residual was also weak (0.103).⁴⁹

Due to the low sample size, standardized model results showed that few of the questions under each construct had estimates above the goal of

⁴⁸ Schraw and Dennison, “Assessing Metacognitive Awareness,” 464.

⁴⁹ Each of these statistical methods attempts to determine how well a model fits the population. The chi-squared test would ideally have a p value of $>.05$ if the model was a good fit. However, the test relies on a larger sample size, which was not present in this study. A good model fit for CFI and TLI would be $>.95$, with larger numbers of variables and lower sample sizes tending to drive the values down. A good model fit for the standardized root mean square residual would be $<.08$. The root mean square error of approximation is another measure of model fit and is similarly biased against complex models with a high number of variables and a low sample size. A good RMSEA would be $<.06$, but a value of less than $<.1$ is at least worthy of consideration. Ultimately, the complexity of this model as well as the low sample size makes determining model fit difficult, so these values are weak, but understandable. See Dexin Shi, Taehun Lee, and Alberto Maydeu-Olivares, “Understanding the Model Size Effect on SEM Fit Indices,” *Educational and Psychological Measurement* 79, no. 2 (2019): 310–34. <https://doi.org/10.1177/0013164418783530>; and Alberto Maydeu-Olivares, “Maximum Likelihood Estimation of Structural Equation Models for Continuous Data: Standard Errors and Goodness of Fit,” *Structural Equation Modeling: A Multidisciplinary Journal* 24 (2017): 390, <https://doi.org/10.1080/10705511.2016.1269606>.

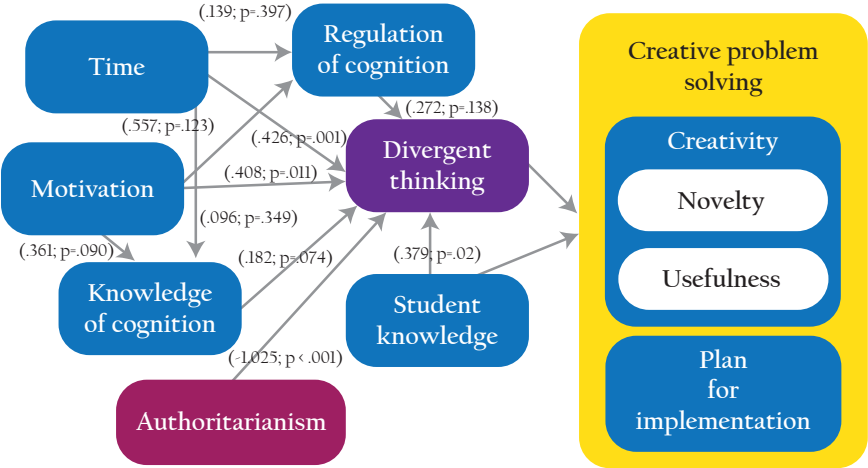
0.7, yet all but two of them had two-tailed p-values that were significant at the 90 percent confidence level. One question within the metacognition (regulation of cognition) construct had a very high p-value ($p=.601$), so it was dropped from the model. Only one other question—within the motivation construct—had an insignificant p-value ($p=.147$), but it was retained in the model due to the small sample size and theoretical fit with the construct. Another question was removed from the student knowledge construct because it was not a good theoretical fit with the construct.

Following the above modifications, root mean square error of approximation improved to 0.082 ($p < .01$). The CFI (0.819) and TLI (0.790) also improved, but it remained weak. The Standardized Root Mean Square Residual also improved (0.098), though it remained weak. The new chi-squared test of model fit ($X^2(302, N = 66) = 436.34, p < .0001$) remained significant. Though many of these values are weak, the model featured more parameters than sample size. These values suggest potential for further research with larger sample sizes.

Confirmatory factor analysis was also conducted on the faculty survey, using factors for divergent thinking, authoritarianism, student knowledge, time and opportunity, motivation, and course structure. Overall, the model fit was poor with a root mean square error of approximation of 0.17 ($p < .001$). The CFI (.439) and TLI (.332) were weak. The chi-squared test of model fit ($X(194, N=31) = 368.056, p < .001$) was significant.

According to the standardized model results, all four of the student time and opportunity items were positively and significantly related to the factor ($p < .1$). Two of the four items had estimates above the goal of 0.7. Additionally, all of the divergent thinking items were significantly ($p < .05$) related to the divergent thinking factor, but one was negatively related. This was the item that asked faculty if they place more emphasis on creativity than content in grading. Two of the five had estimates above 0.7. For the student knowledge factor, two of three items were significantly related to the factor ($p < .01$) while the third was not significant ($p = .169$). Only one of the items had an estimate above 0.7. Standardized results for the other factors showed that less than one-half of the items were significantly related to their hypothesized factors. Of note, the fac-

Figure 2. Conceptual framework revisited



Source: courtesy of the author, adapted by MCUP.

tor for student time and opportunity was significantly and positively related to the factors for divergent thinking (.622; p = .006) and student knowledge (.649; p = .041), and significantly and negatively related to the factor for authoritarianism (-.435; p = .041).

Conceptual Framework Revisited

An analysis of covariance for each of the factors originally introduced in the conceptual framework shows significant covariance among several of the factors on the student survey. The factor for metacognition has been divided into its component parts (knowledge of cognition and regulation of cognition). The estimate and two-tailed p-value for each hypothesized relationship between factors is illustrated in figure 2. Many of the hypothesized relationships proved to be significant. The largest and most significant relationships included the relationship between divergent thinking and time and opportunity (positive) as well as divergent thinking and authoritarianism (negative).

Table 3. Student responses to prompt asking if they improved their creative problem-solving skills as a result of attending CSC

Response	N	%
Strongly agree	20	25.6%
Agree	35	44.9%
Neutral	20	25.6%
Disagree	2	2.6%
Strongly disagree	1	1.3%

Source: courtesy of the author.

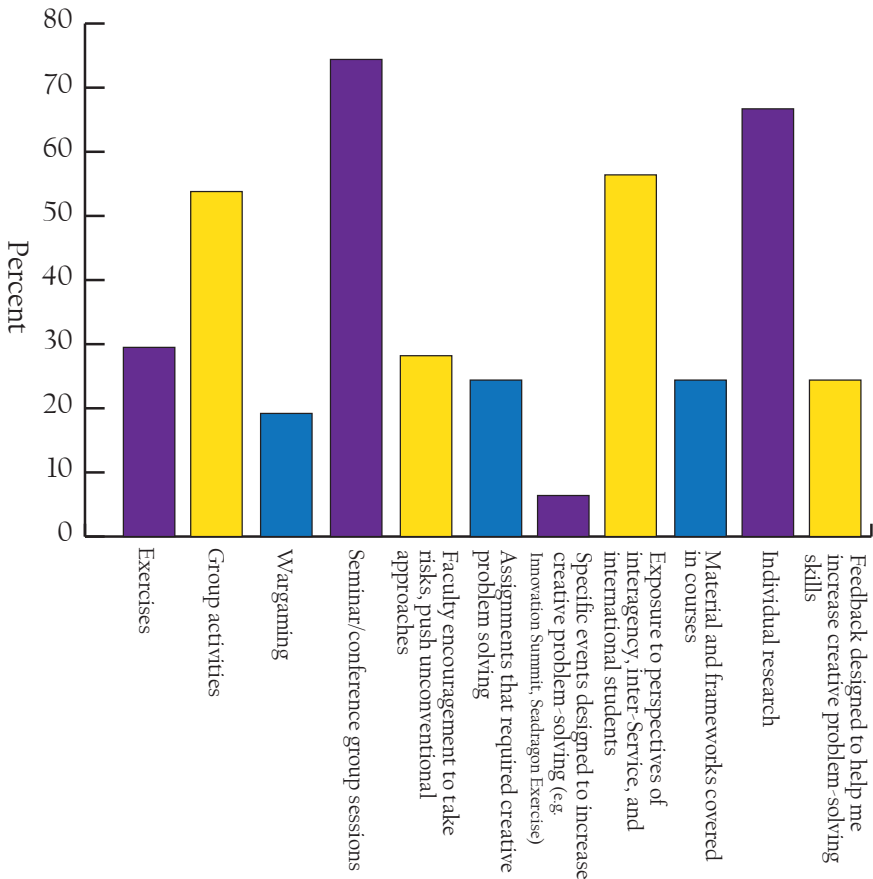
Summary of Results and Key Findings

Results and findings are reported here, beginning with the analysis of secondary data, both quantitative and qualitative. Secondary data included annual student surveys, assessments of student creative problem-solving ability, and grades. Original data include publicly available master’s theses and the surveys specifically developed for this study (featuring both qualitative and quantitative data). Additionally, for student surveys in which the students voluntarily disclosed their names, survey data was compared to faculty assessments of student work via multiple linear regression.

Secondary Data: Prior Student Annual Surveys

For the university’s standard annual survey from year one of the study, the researcher obtained secondary data in the form of responses to questions related to creative problem solving. In the closed-ended questions, students were asked to indicate on a five-point Likert scale from strongly agree to strongly disagree if their creative problem-solving skills improved as a result of attending CSC (table 3). The responses were normally distributed and centered on “agree” with a very small tail toward “disagree.” In total, 78 students responded to this question and 70.5 percent of students indicated that they agreed or strongly agreed that their creative problem-solving skills improved as a result of attending CSC. Another 25.6 percent of students were “neutral.” Only three students indicated that they disagreed or strongly disagreed.

Figure 3. Most significant influences on improvement of creative problem solving (n=78)



Source: courtesy of the author, adapted by MCUP.

After answering this question, students had the opportunity to select what were the most significant influences on improving their creative problem-solving skills (figure 3). Students could select as many influences as they wanted. Two influences emerged most commonly. seminar/conference group sessions (74 percent) and individual research (67 percent). Two other influences exceeded 50 percent: group activities (54 percent) and exposure to perspectives of interagency, inter-Service, and international students (56 percent).

Subsequently, students were given the opportunity to provide additional open-ended comments. Two positive themes emerged. The most prominent theme was praise for the classroom environment, including the faculty and peers. For instance, when given the opportunity to describe significant influences on their creativity, one student said: “I thought the experiences of the [military faculty] MILFACs [name redacted] and my fellow students based upon their past experiences was very valuable.” Another positive theme was the opportunity to take electives (even though only 2 of approximately 41 credits were electives). Negative themes that emerged were that wargames and exercises were not helpful and that the focus on creating conventional planning products detracted from the ability to think creatively, both in time and structural constraints. One student, conversely, strongly recommended more wargaming.

Students were next given the opportunity to provide open-ended responses concerning barriers to developing creative problem-solving skills at CSC. Many students felt that assignments and feedback incentivized conventional thinking rather than unconventional. Some students complained that faculty did not support students who provided diverse opinions, “challenged the status quo,” or engaged in “out-of-the-box thinking.” Another key theme was the lack of time to think creatively about what the students were learning. One student said,

The fire hose effect of pushing material onto students without affording them the opportunity to stop, think critically, and then apply what you learned. . . . If you force 80 plus pages of reading a night, which takes time to read, comprehend, and digest, students don’t have time to think critically about what they learned and apply it to further studies. It turns into a one-night digestion where you regurgitate it the next day and move onto the next 80 pages of reading.

Secondary Data Analysis of Student Creative Problem-Solving Ability

For the first year of the study, the researcher obtained more detailed information about the assessments used for the university’s quality enhancement plan, including the specific assignments assessed and the differences

between the ratings of the military and civilian faculty at CSC. For the fall semester, the rated assignment was an analytic essay on grand strategy and rogue states that was part of the National Security Affairs and the International System Course. The civilian faculty were the primary instructors for this course and an essay was due at the end of the semester. The prompt asked students to either assess the current state of U.S. grand strategy or to craft a new U.S. policy toward Iran or North Korea.

For this essay, civilian faculty scored 120 papers using the rubric, while military faculty scored 150 papers. Of these papers, 79 were scored by both a civilian faculty member and a military faculty member. Overall, 40 percent of papers were assessed to be creative or transformative. For the 79 papers assessed by both a military and a civilian faculty member, the average score as assessed by civilian faculty was 2.3, while the average score as assessed by military faculty was 2.7. Hence, the civilian faculty rated student creativity, on average, almost one-half point lower than the military faculty on this assignment. Moreover, the civilian faculty found only 41 percent of these papers to be creative or transformative, while the military faculty found 63 percent of the same papers to be creative or transformative.

There were 23 papers for which the average score differed between raters by one point or more. On all but four of these instances, the civilian faculty member's rating was lower than the military faculty member's. For a further 16 papers, the average scores differed by less than one point but more than a one-half a point. For only three of these papers, the civilian faculty member's rating was higher than the military faculty member's rating. Overall, taking the average scores for each paper, there was more than half a point difference between the military and civilian faculty ratings on 49 percent of the papers, demonstrating substantial disagreement between civilian and military faculty members.

For the spring semester, the essays evaluated were part of Leadership in the Profession of Arms II, for which the military faculty are the primary instructors. The prompt asked students how the U.S. military might be wrong and how to fix it. Overall, 45 percent of papers were assessed to be creative or transformative, but substantial differences emerged when

Table 4. For year one of the study, percent of papers rated creative or transformative (among papers rated by two faculty members)

	N	Military faculty	Civilian faculty
National security (Fall)	79	63%	41%
Leadership (Spring)	14	57%	92%

Source: courtesy of the author.

comparing the scores of the military faculty raters to the scores of the civilian faculty raters. These differences were the inverse of the previous essay. Of the 46 papers civilian faculty rated, they found that 76 percent were creative or transformative. On the other hand, of the 54 papers that military faculty rated, they found that only 24 percent were creative or transformative. Only 14 papers were assessed by both a civilian faculty member and a military faculty member. Within this set of papers, the average rating from the civilian faculty was 3.1 while the average rating from the military faculty was 2.6, exactly one-half a point lower. Overall, for one-half of the papers rated by two faculty members, the average scores were more than half a point different between raters. Moreover, of these 14 papers, civilian faculty rated 92 percent of them as creative or transformative, while military faculty rated only 57 percent of the same papers as creative or transformative.

Considering both semesters together, an interesting pattern emerges for those papers evaluated by both a civilian faculty member and a military faculty member. The civilian faculty members, on average, rated the same paper lower than military faculty for a subject in their general areas of academic expertise (national security affairs and the international system). Meanwhile, for a subject in the military faculty’s general experiential area of expertise (leadership), the military faculty rated students lower, on average, than the civilian faculty. So, in both cases, the area of faculty expertise seems to have negatively influenced the perception of student creativity in that area (table 4).

Year Two Student Assessments and Grades

Unlike academic year one, CSC used a staff ride exercise to evaluate student creativity in academic year two.⁵⁰ For the first semester, students went to Chancellorsville, Virginia, for the exercise on 9–10 November. Faculty conducted 211 evaluations of creativity with 145 students based on oral presentations at the staff ride. For those students who were evaluated twice, their scores were averaged. Fifty-two percent of students were adjudged to have produced creative or transformative work on the staff ride. Table 5 shows descriptive statistics for their scores for each element of the AAC&U *Creative Thinking VALUE Rubric*.⁵¹

Table 5 also shows student grades on two essays that were designed to evaluate student achievement on learning outcomes related to critical and creative thinking. The first essay was the future war and innovation essay from the 3100 Evolution of Modern War Course. This essay asked students to describe lessons that today’s national security community might learn from decision making between World Wars I and II. It was submitted in December. Another essay on great power competition asked students to propose changes to U.S. strategy or policy. This essay was due in November and was part of the 4100 National Security Affairs and the International System Course. Table 5 below shows the average grades for each of these essays as well as the first semester GPA.

Quantitative Survey Results

As previously discussed, surveys for faculty and students were developed to measure factors related to creativity in PME. They reflect faculty and student thoughts at the end of the first semester of the year and they occurred within a reasonably close time to the QEP assessment, essay grades, and first semester GPA described above. These surveys used a five-point Likert scale (1= strongly disagree to 5= strongly agree) for questions in the following factors: divergent thinking, authoritarianism, student

⁵⁰ Bradford Wineman, Christopher S. Stowe, and Paul D. Gelpi, “The Staff Ride: A Model for the 21st Century,” *Marine Corps Gazette* 102, no. 7 (July 2018): 70–75.

⁵¹ *Creative Thinking VALUE Rubric*, 2.

Table 5. Descriptive statistics for grade point average (GPA), grades, and QEP assessments (n = 209)

	Grades	QEP assessment								
	1st semester GPA	3100 essay grade	4100 essay grade	Acquire strategies/skills	Embrace contradictions	Connecting and synthesizing	Innovative thinking	Solve problems	Take risks	Total
Mean	3.67	90.99%	90.05%	2.19	2.31	2.38	2.31	2.51	2.31	2.34
Stdev	0.25	3.45%	3.70%	0.77	0.91	0.91	0.92	0.87	0.87	0.88
Min	2.7	82%	80%	0	0	0	0	0	0	
Max	4.0	98%	98.25%	4	4	4	4	4	4	

Note: Goal for QEP is that 80 percent of students will score a 2.5 or higher.
Source: courtesy of the author.

knowledge, time, motivation, and metacognition. Due to the relatively small population, surveys were sent to all CSC students and to any faculty member who taught a CSC seminar (including electives, but excluding contractors) or mentored CSC student theses. A total of 66 students and 31 faculty members responded to the surveys. In the section that follows, each factor will be examined in greater detail.

Factors

Each survey had three to five questions about each factor, except that the factors related to student metacognition only appeared on the student survey. For those factors present in both the student and the faculty surveys, most questions on the student survey had parallel versions on the faculty survey. Table 6 summarizes faculty and student survey results for each construct.

Overall, faculty and students had remarkably similar scores for each of the factors that were included in both surveys. Unexpectedly, given the relative sample sizes, standard deviations for most of the factors were higher in the student survey than in the faculty survey. This finding suggests greater variability amongst the student body concerning perceptions of the CSC environment.

As would be hoped, among factors on both surveys, students and fac-

Table 6. Average and standard deviation for each factor

Factor	Faculty (n=31)		Students (n=66)	
	Mean	Std Dev	Mean	Std Dev
Divergent thinking	3.97	.59	3.85	.5
Authoritarianism	2.26	.64	2.39	.97
Student knowledge	3.33	.81	3.81	1.15
Time and opportunity	2.59	.83	2.32	1.00
Motivation	3.46	.78	3.46	1.24
Metacognition (knowledge of cognition)	n/a	n/a	4.20	.61
Metacognition (regulation of cognition)	n/a	n/a	4.16	.83

Source: courtesy of the author.

ulty reported the highest average scores for divergent thinking, and these scores had the lowest standard deviation. Based on these responses, there seems to be general agreement that divergent thinking is encouraged and practiced at CSC. Both groups also averaged below the midpoint of the Likert scale for authoritarianism, which is a good sign. On the negative side, for students, the factor for time and opportunity had the lowest average score of any of the constructs. The scores on this factor indicate that students, on average, did not think that they had enough time to think creatively about the content of the curriculum. Moreover, the faculty largely agreed that students do not have enough time to think creatively.

Both faculty and students returned relatively high average scores for motivation and student knowledge. These student knowledge scores comprise knowledge of the content as well as knowledge of research methodology. One possible reason that both faculty and students report a lack of time is that there is so much content to cover. Yet, the results of these surveys indicate that the students and faculty feel relatively comfortable with student knowledge.

Finally, for the two factors related to metacognition, students reported their highest average scores. Students seem comfortable with their

understanding of how to process information (knowledge of cognition) and their ability to control their own thinking while solving problems (regulation of cognition). Taken together, the three factors hypothesized to positively relate to creative thinking skills (divergent thinking, knowledge of cognition, and regulation of cognition) had the three highest average scores and the three lowest standard deviations on the student survey. Students were very confident in their thinking skills.

Relationship between Factors and Creativity Outcomes

Because some students provided their names, their survey data for each of the above factors could be compared to their first semester GPA, essay grades, and QEP assessments (if they were in the CSC QEP sample). Students who provided their names on the student survey (n=27) had average scores on these essays that were within one percentage point of the overall average for the 3100 essay and one-tenth of a percentage point of the overall average for the 4100 essay. Similarly, the percentage of students adjudged to be creative or better (50 percent) who provided their names on the student survey discussed below and were evaluated on the QEP (n=16) was very similar to the total percentage of students in academic year two who were adjudged to be creative or better (52 percent). Finally, the average first semester GPA for students who provided their names on the student survey in December was within .05 points of the overall average GPA. Overall, there was substantial similarity in outcome measures (GPA, essay grades, and QEP performance) between students who submitted their names on the student survey discussed below and those who did not submit their names.

Using multiple linear regressions ($Y = b_0 + b_1X_1 + b_2X_2 + b_3X_3 + b_4X_4 + b_5X_5 + b_6X_6 + b_7X_7$), where Y reflected the results on each of these assessments, table 7 shows t-scores for each factor on each assessment. Because only 27 students provided their names, and only 16 of those students were in the QEP sample, it was difficult to establish significant relationships between these assessments of student learning and the results of the student survey, even at the 90-percent confidence level. Detailed regression analysis on such small sample sizes is not normally advised, but this was the only method available for making the connection between survey results and

Table 7. T-scores for each factor with grades and QEP assessments as dependent variables (n= 27 for 3100, 4100, GPA; n= 16 for QEP)

	3100	4100	GPA	QEP
Divergent thinking	-1.227	-0.375	-1.065	0.65
Authoritarianism	-1.488	-1.334	-1.735*	-2.085*
Student knowledge	1.115	-0.567	0.936	-0.309
Time and opportunity	-2.433**	-1.737*	-1.256	-1.762
Motivation	0.167	0.527	-0.553	-0.007
Metacognition (knowledge of cognition)	0.275	-0.753	-0.238	0.946
Metacognition (regulation of cognition)	-0.390	-0.648	-1.42	-2.324**

*p < .1; ** p < .05
Source: courtesy of the author.

assessed creative problem-solving ability—a key element of the study. Results reported below suggest the need for further research.⁵²

The analysis in table 7 shows some expected and unexpected results. As expected, authoritarianism was significantly and negatively related to QEP scores. That this relationship extended to GPA was also an interesting finding, suggesting that creative thinking might not be a barrier to academic success—in spite of some student survey comments that indicated that some faculty did not reward creative thinking in their grading. Moreover, the negative relationship between time and opportunity and grades on the 3100 and 4100 essays was surprising. This result would seem to suggest that as students agreed more strongly that they had enough time to think creatively, they actually scored significantly worse on these essays. Similarly, as students reported higher scores on regulation of cognition, they scored significantly worse on the quality enhancement plan’s assessment of creativity. Moreover, there were no other significant effects of metacognition and no significant effects of divergent thinking on these assessment scores.

⁵² For detailed regression analysis tables, see Sparks, “Fostering Graduate Student Creative Problem Solving,” 207–10.

Theses

As discussed previously, students and faculty rated student CSC content knowledge relatively highly. The content knowledge measure included the idea of research literacy. One way to break out research literacy is to analyze how many students conducted empirical research as part of their final theses. Using a publicly available database of student theses, this author selected a random sample of 30 theses to determine how many of the theses use at least basic empirical research methodology.⁵³ Basic empirical research methodology is defined as a study conducted by the student that goes beyond a literature review and research question, where the population to be studied, the methodology, the findings, and a discussion are reported.⁵⁴ For the purposes of the analysis, historical case studies were not considered to be empirical research. Of course, there is nothing wrong with theses that do not use empirical research methodology, but the number and quality of theses in which the students conduct empirical research is a potential indicator for their understanding of such research.

In the sample of 30 theses reviewed, all from year one of the study, 4 were identified in which the student conducted basic empirical research, though none of them could be said to be experimental or quasi-experimental.⁵⁵ Of those four theses, two won awards. Of the remaining 26 theses reviewed, 1 other thesis won an award and the thesis was a historical case study. For comparison purposes, a total of 15 students across CSC in year one of the study received writing awards. Some of these were for other papers, not the thesis. This suggests that faculty highly value empirical research methodology when identifying exceptional work but that students rarely conduct it.

Qualitative Survey Themes

Both faculty and students had the opportunity to respond to option-

⁵³ "CSC Student Papers AY 2021," CSC Student Papers Collection, MCU Student Papers, Marine Corps Research Library, accessed 4 March 2024.

⁵⁴ "Empirical Research in the Social Sciences and Education," Library Guides, Penn State University Libraries, accessed 4 March 2024.

⁵⁵ William Shadish, Thomas D. Cook, and Donald Thomas Campbell, *Experimental and Quasi-experimental Designs for Generalized Causal Inference* (Boston, MA: Houghton Mifflin, 2002).

al open-ended questions at the end of the surveys that were developed specifically for this study and administered in year two. The open-ended questions concerned barriers to creativity, bridging the potential divide between military hierarchy and divergent thinking, and strategies for creative thinking in PME. A total of 43 students and 17 faculty provided at least one response to an open-ended question. There was not a significant difference in average responses for each factor discussed above between respondents who provided open ended responses and those who did not.

Using first-cycle emergent coding and second-cycle theoretical coding, a few key themes emerged from the qualitative analysis of the open-ended survey responses.⁵⁶ Thematic analysis resulted in 6 themes and 21 subthemes. Table 8 lists these themes, subthemes, and example codes within each subtheme.

Time

Time was the key theme for the open-ended survey responses for both faculty and students. Time to think and the amount of reading were the two most often mentioned codes for both faculty and students. The clear implication with both groups was that the amount of reading was too great and it impeded students' ability to dedicate time to thinking. For instance, one student described the amount of reading required each day as "ridiculous," while another student stated that "students would benefit from less required reading per night . . . in order to provide time to adequately digest the subject matter and to facilitate critical thinking."

Several of the faculty comments indicate agreement about lack of time to think, with one faculty member attributing the problem to thinking in the government world:

People need space and time to think divergently about topics. [CSC] takes the opposite approach; CSC believes that more activities, assignments, lectures, readings, etc. equals more rigor. That is absolutely not the case. Unfor-

⁵⁶ Victoria Elliott, "Thinking about the Coding Process in Qualitative Data Analysis," *Qualitative Report* 23, no. 11 (2018): 2855, <https://doi.org/10.46743/2160-3715/2018.3560>; and Johnny Saldaña, *The Coding Manual for Qualitative Researchers* (Los Angeles, CA: Sage Publications, 2016), 250.

Table 8. Themes from open-ended survey responses

Theme	Subthemes	Example codes
Operations	Organization	Vertical integration
	Schedule	Flexibility
	Content	Historical case studies
	Context	Job applicability
	Process	Teach creativity
	Quality	Depth of content
Divergence	Encouraging	Set the tone
	Lack of	Words versus actions
	Safety to	Outcasts
Faculty	Authority	Close-mindedness
	Quality	Variation
Locus of problem	Faculty	Military faculty
	Students	Student knowledge
Metacognition	Regulation of cognition	Planning learning
	Thinking techniques	Red teaming
Time	To discuss	Coverage of readings
	To read	Page requirement
	To think	Busywork
	To work	Requirements
	To write	Structured writing time
	Time management	Time dictates quality

Source: courtesy of the author.

Unfortunately, in the government world, there is a notion that if the time is not organized with some sort of activity, then employees must be goofing off. No—we need time to read, write, and think about this stuff.

Especially in the faculty comments, there were complaints about the amount of time required for assignments, and not just for reading. Moreover, students indicated that they actively managed their time, putting

more or less effort into each assignment, depending on the amount of time they felt that they had available to complete the assignment.

Content

Students and faculty also complained about the content of the curriculum. There were calls to reduce historical content and then other calls to increase the historical content. Additionally, there were calls to focus the content of the courses on practical problems related to the military professional and other calls to diversify the curriculum to incorporate more lessons from the business world and other areas. Both faculty and students seemed to agree that the curriculum should be more focused and have greater depth. For instance, one faculty member suggests that the curriculum should “move from covering the waterfront in what feels like a replay of college/BA/BS, purposefully focus on fewer but key themes to do a deeper dive.” Other faculty complain that some students do not “have sufficient academic preparation and/or command of written English to successfully complete a graduate-level program.” Ironically, however, there was little agreement about where to focus the curriculum. Moreover, students and faculty also called for a greater diversity of viewpoints in the readings at the same time as they called for fewer pages of reading.

Authoritarianism versus Divergent Thinking

There were a wide variety of opinions about the environment for divergent thinking at CSC as well as the existence of authoritarianism that negatively impacts creativity. Some students and faculty spoke highly of the academic freedom they possess, while others decried the inability to discuss controversial topics in the seminar group. Furthermore, when students reported a problem with authoritarianism, they did not always agree on where the problem laid. Some students attributed the problem to military faculty members who did not understand the academic environment and sought only to indoctrinate students with their own views. Other students questioned the assumption that military hierarchy was the problem and instead pointed to civilian faculty as creating an environment that was hostile to different ways of thinking. Still other students

attributed hostility to their fellow students, who might have lacked the maturity necessary to discuss controversial topics.

Two key suggestions for creating an environment of academic freedom, psychological safety, or divergent thinking emerged from the faculty and student surveys. First, several respondents suggested that students and faculty should not wear uniforms with rank during the seminar discussions. The argument was that some students will self-edit to avoid the risk of irritating senior officers and that the wearing of rank in the classroom is a constant reminder of seniority. One student, on the other hand, specifically stated that the removal of rank “is a childish and cowardly away to approach having open and professional discussions.”

The second key recommendation for encouraging divergent thinking emerged primarily from the faculty survey where faculty members recommended setting a tone early on in the year that respects divergent opinions. One way faculty recommended this be done was to praise students who raised divergent opinions and challenged the status quo while never “responding emotionally or defensively to a student who is trying to think critically.” Of course, doing this is difficult and some students complained that initial verbal support for divergent thinking did not translate into the actual classroom experience. However, both students and faculty seemed to recognize that there was a difficult balance to maintain between respect for military authority and encouragement of outside the box thinking.

Thinking

Finally, the student survey asked students to give some examples of ways that they manage their thinking when they are solving problems. Here again, time management was a key theme, with one student spelling out a long, sarcastic response about the process of developing ideas and then concluding that they really just “think about the problem for a bit, usually reference some things I’ve read or remember reading. Bounce some ideas off a couple people who are smarter than me. Sleep on it. Finalize the ideas in the shower and on my commute.”

In addition to the theme of time management, some students also mentioned specific thinking techniques like playing the devil’s advocate,

cost-benefit analysis, crowd-sourcing, brainstorming, writing the problem down and returning to it later, or referring to specific structured analytical techniques. Students also mentioned that they plan how they are going to solve a problem, try to monitor their thinking as they are solving the problem, or seek feedback on their ideas. Specific sources of feedback include their fellow students, mentors, and faculty. Finally, more than one student mentioned taking time off to think about something else and then returning to the problem later.

Discussion

The research questions asked about faculty and student perceptions of factors related to creative thinking. They also asked how students' perception of these factors related to their performance on assessments designed to encourage creative problem solving. Additionally, this study aimed to identify how faculty and students described the environment for creative problem solving at MCU. This discussion will consolidate findings for each of the factors across the various data sources.

Divergent Thinking

The quantitative results show that students and faculty generally had a high view of divergent thinking at CSC, and divergent thinking was significantly and negatively correlated to authoritarianism on the student surveys. Given the chance to express their thoughts, some students were highly complimentary of specific faculty members. At the same time, open-ended responses tended to focus on problems. However, the locus of the problem was often not necessarily the faculty. Moreover, problems and attitudes toward divergent thinking and creativity in general often varied by the individual faculty member, such that some faculty were perceived to be exceptionally good and others were perceived to be exceptionally bad.

Furthermore, opinions about student divergent thinking often varied by faculty member. For instance, the fact that faculty opinions of student creativity tended to decrease when students were working in that faculty member's area of expertise is an important finding. Additionally, students sometimes reported a lack of standardization in grading or that grading

did not reward creativity. Disappointingly, divergent thinking was not a significant predictive factor for scores or grades on any of the assessments or on GPA. Perhaps this is partially explained by the fact that divergent thinking ratings were so high on the scale and had low variability.

Authoritarianism

As with divergent thinking, student perceptions about authoritarianism were highly faculty dependent. Meanwhile, faculty and students both generally recognized the need to balance military hierarchy with psychological safety in the classroom, to greater or lesser degrees of success. The fact that authoritarianism and divergent thinking were significantly and negatively correlated on the student surveys suggests that any curricular intervention aimed at increasing creativity should consider the relationship between divergent thinking and authoritarianism. In addition to being significantly and negatively related to divergent thinking on the student surveys, authoritarianism was also significantly and negatively correlated to GPA and the practical assessment of creative thinking at the battlefield staff ride.

Student Knowledge

Aside from a few negative comments about select students, faculty and students both had a relatively high opinion of student knowledge. Students and faculty both generally felt that student knowledge of research methodology was high even though a review of the student theses shows that students did not often use empirical research methodology. Additionally, some students and faculty wanted to allow for deeper dives into various content areas, sometimes to increase student knowledge as a foundation for creative thinking and sometimes to allow students to examine topics of greater interest to them.

Time and Opportunity

The most important finding in all of the qualitative data as well as the descriptive quantitative data is that there was a perception that students did not have enough time to think creatively. Many respondents, especially students, attributed this lack of time to an overwhelming amount

of reading or other work required to ensure students have good content knowledge. Given that student knowledge seemed to meet expectations, one might conclude—as many respondents did—that the amount of reading could be reduced. At the same time, perhaps a reduction in the amount of reading might reduce student knowledge to unacceptable levels. Similarly, for both faculty and student surveys, the factor for time and opportunity was significantly and positively correlated to the factor for divergent thinking.

The perception that students did not have enough time to think does not bear out in the analysis of scores on the student products examined for creativity. In fact, confusingly, scores for each of the examined analytic essays were significantly and negatively related to the factor of time and opportunity on the student survey. This indicates that students who thought they did not have enough time to think scored higher than those who thought they did have enough time. Perhaps these findings rule out the perception that students who perform poorly use lack of time as an excuse. On the contrary, it seems that the students who perform the best also perceive that they do not have enough time, potentially suggesting that they are putting extra effort into the work. Therefore, the relationship between time and opportunity and divergent thinking deserves further examination.

Motivation

Motivation was not an important theme of the qualitative results as neither students nor faculty mentioned it. Moreover, though student motivation was significantly and positively correlated to divergent thinking and knowledge of cognition on the student surveys, regulation of cognition was not significantly correlated with motivation. These results suggest that student motivation is not one of the most important factors for creativity in this context.

Metacognition

Faculty were not specifically asked about student metacognition, and the idea did not emerge as a theme from the faculty qualitative survey results. Students were asked about metacognition and had a very high degree of

confidence in their metacognitive ability. The survey had two subfactors, knowledge of cognition and regulation of cognition, both of which returned high scores on the student survey. Oddly, one factor (knowledge of cognition) was significantly and negatively correlated to divergent thinking, while regulation of cognition was not significantly related to time, motivation, or divergent thinking. Moreover, another metacognition factor (regulation of cognition) was significantly and negatively related to scores on the evaluation of creativity in the battlefield staff ride, but it was not significantly related to any of the other assessments or GPA. As with the confusing results related to time and opportunity, this result was also confusing in that it suggests that high regulation of cognition led to a lower creativity score. However, the items used to identify these factors were only a small subset of the items from the larger metacognitive awareness inventory (MAI).⁵⁷ The larger MAI was not used for this research due to its size and the time required to complete the survey. Perhaps a deeper analysis using the full MAI might return different results.

When asked to describe tactics they used to solve problems, students often described methods for regulating their cognition, especially in preparation and during the process of problem solving. However, they were also clear that time dictated how much planning and monitoring they could support. Given the above findings related to time, it seems that these two factors might go hand in hand. Perhaps giving students time to think might be insufficient to increase their creativity. Possibly students also need support to use that time well. Likewise, perhaps students do have good metacognitive ability but do not have time to exercise it. Indeed, one student complained that they did not get the prompts for the battlefield staff ride in time to prepare a good response. As with the factor related to time and opportunity, there is room for a more thorough examination of the relationship between metacognition, divergent thinking, and creative problem-solving at MCU.

Limitations

Two major limitations should be noted for this assessment. Of greatest

⁵⁷ Schraw and Dennison, "Assessing Metacognitive Awareness," 472-74.

importance, during year two, the sample sizes for the student survey ($N = 66$) and faculty survey ($N = 31$) were low, raising the possibility of type 2 error because of the low power of the sample. To mitigate this concern, significance was set at $p < .1$ rather than the common $p < .05$ standard, but this increases the possibility of type 1 error.

Secondly, this study used data primarily from the COVID-influenced academic years, which might have an unknown effect on the overall data, making these academic years less likely to be generalizable to non-COVID influenced years. In both academic years, students mostly continued to meet in person with some events (e.g., lectures in year one) moved into a virtual format. Additionally, some field trips were restricted in year one and students and faculty were periodically required to meet virtually in seminar due to COVID-exposure. COVID-19 also may have had unknown psychological or emotional impacts, which might have influenced student learning and creativity.

Conclusion

This study uncovered some important findings about the environment for creative problem solving at CSC. Students and faculty generally perceived divergent thinking to be high and authoritarianism to be low. However, it was unclear how divergent thinking related to assessments of creative problem solving. Additionally, some of the qualitative data show that perceptions of authoritarianism are highly dependent upon individual faculty members.

The perception that time was not sufficient for creative problem solving was the single most widely shared finding in the qualitative data and the descriptive quantitative data. However, the relationship between student assessment data and perceptions that there was not enough time for creative problem solving was ambiguous or even sometimes significantly and negatively related. Perhaps students who felt that they did not have enough time were also the students who put forth the most effort. Their grades reflect their effort, even though they would have liked to have more time to think creatively.

Similarly, the findings concerning metacognition deviated from the expected. Though students generally perceived their metacognition to be

high, the metacognition scores were ambiguously related to student assessment results. Additionally, metacognition scores were not related to perceptions of time and opportunity. For future work, the relationship between metacognition and time and opportunity deserves more inquiry with a more robust metacognition measure.⁵⁸

⁵⁸ For follow-on research, see Sparks, “Fostering Graduate Student Creative Problem Solving,” 112–72.



CHAPTER 3 ●

Creative Community of Learners—the Auxiliary Engine to Andragogical Approaches

Case Studies from the French War College

Emilie Cleret and Jerome Collin

École Militaire was built in the mid-eighteenth century. Within its premises, visitors will find *École de Guerre*, the French War College. It is located opposite another famous landmark, the Eiffel Tower, erected in the late nineteenth century.

Creative communities of learners have become essential to talent development in certain key areas of senior professional military education in France, such as increasingly the level of the officers' interoperability with their counterparts from allied countries such as the United States, Canada, the United Kingdom, or other European countries.

The French Language Department and the English studies curriculum at the French War College share similar goals (i.e., increasing the interoperability of people for effective communication in coalitions). This means that similar challenges are also shared. To overcome some of those challenges, both departments have focused on the idea of creating creative communities of learners. From that experience, some interesting conclusions were drawn about andragogical approaches.

This chapter first focuses on the experience initiated by the French Language Department faculty, which blended students living outside of their home country into an immersive learning environment with students in a domestic learning framework (i.e., learning French within their home country). The goal of this experience was to help improve authen-

tic and effective communication in the workplace. The second part of this chapter describes a different type of community of learners working within the framework of the English studies' international curriculum. These communities were created to trigger a positive clash of skills to support the learners' awareness of how much they could learn from each other, and through this trading of skills, an increased interoperability of people was able to emerge.

Communities of French Language Learners

The way the French Language Department approached its first initiative to create a community of learners in the spring of 2022 was to build on the basic principle that there are two different methods for learning a foreign language.¹ The first is to study the foreign language in one's own country, with a language teacher from one's own culture, who was born speaking the same language as all the learners. We refer to this as domestic learning.² The second consists of going to a country where the language one is learning is spoken, attending classes delivered by teachers who live in that country, and speaking the language one is studying on a daily basis. We refer to this as immersive learning.³

Each method has its advantages and disadvantages. Domestic learning enables the learner to grasp the basics of the language in a comfortable and familiar setting, since the course takes place in the learner's own country, with teachers of the learner's own nationality. The teaching methods are those to which the learner is accustomed, and cultural references are introduced and made explicit with the learner's own cognitive filters. In the case of domestic learning, groups are generally linguistically homogeneous: the phonetic and grammatical approach is favored by the

¹ This initiative was developed during two educational visits by Madli Kütt (Estonian Military Academy, Tartu) in 2022 (28 March–1 April) and in 2023 (29 May–2 June).

² In French, we refer to *apprentissage linguistique à domicile* or *formel*, which means learning at home, in school, or at university in the learner's country and/or without contact with a population using the language studied on a daily basis.

³ In French, we refer to *apprentissage linguistique en immersion*, which means learning carried out in a region where the language studied is used daily, with a pedagogy exploiting the presence in the study environment of local speakers and the written language.

possibility of using comparative phonology, or elements of comparative grammar. Learning progression can be calibrated according to a fine-tuned hierarchy of difficulties identified in the transition from the logic of one particular language to that of another.

Immersive learning, conversely, brings the learner into contact with the authentic language and its associated culture. The language is heard not only in the learning environment, but also in the public sphere. The learner's need to use the language and to put learning into practice is a daily occurrence, which is a strong motivating factor and enables the learner to apply what they have learned immediately in everyday life, by coming into contact with a society that expresses itself in the target language. It also gives learners the means to test their ability to make themselves understood and to develop strategies to improve their efficiency in real time.

However, both of these learning methods also have drawbacks.⁴ Domestic learning is generally based on purely academic methods, which are often far removed from the living practice of the language and therefore artificial, and does little to prepare learners for the sociolinguistic reality they would have to face when they find themselves having to communicate in countries where the language studied is spoken on a daily basis. This learning is divided into dedicated periods, sometimes very short (less than an hour) and spaced out over time (a few days a week). The rest of the time, learners may be busy with other preoccupations and in an environment that distances them from their language learning. The sudden discovery of the country where learners will be practicing the lan-

⁴ This reflection on the limits of the two modes of learning when it comes to learning a foreign language owes much to the difference between unconscious (immersive) learning and intentional (domestic) learning. Lev S. Vygotsky writes about classical learning of a foreign language that “the child assimilates his mother tongue unconsciously and unintentionally, whereas learning a foreign language begins with awareness and the existence of an intention. This is why we can say that the development of the mother tongue is from the bottom up, while that of the foreign language is from the top down. If the development of the mother tongue begins with its spontaneous and easy practice and ends with awareness of its verbal forms and their mastery, the development of the foreign language begins with awareness of the language and its voluntary mastery and ends with easy and spontaneous speech.” Lev S. Vygotsky, *Thought and Language*, trans. Alex Kozulin (Cambridge, MA: MIT Press, 1986).

guage can be a major obstacle. During the learning process, learners are often confronted with a fantasized and caricatured vision of the target language's sociolinguistic universe, far removed from the reality they will encounter. This is particularly true in the case of France and Paris, which are presented in textbooks with numerous clichés that may be at odds with the reality and difficulties of everyday life. In the manual *Le nouveau Taxi!*, learners are asked to work on the following model, which they have to adapt to the city of Marseille, from a purely touristic and idyllic angle: "I'm in Lille in the north of France. It's a very pretty town. We have a 50 m² two-bedroom apartment: it's very nice. We go to the movies, bars. . . . We also visit museums."⁵ In the "Alter Ego 4" textbook, in one of the first lessons focused on a so-called French vision of love, there is an illustration to be described and titled, showing a cliché of lovers on a bench with the caption: "Rain of kisses with heartbeat. Persistent shivers at the end of the day."⁶

But immersive learning can also create mental blocks to language use and cultural acceptance in the country of study. The difference between the language learned at school and the language spoken in everyday life can discourage learners. When they leave the classroom, they discover a language in the streets that seems different and much more complex than the one they are learning. They may also find it difficult to adapt themselves to the cultural reality of the country, especially when their own cultural or social cues (dress code, rhythm, behavior) are different. Learners may project onto the language itself the judgment they make about the society they are discovering. Further, immersive learning groups can

⁵ Guy Capelle and Robert Menand, *Le Nouveau Taxi*, vol. 1, *Méthode de français* (Paris: Hachette, 2009), 41: "Chère Pauline, Je suis à Lille, dans le Nord de la France. Cette ville est très jolie. J'habite dans le centre avec une amie espagnole : elle va aussi à l'institut avec moi. Nous avons un appartement de 50m², avec deux chambres : il est très bien. On va au cinéma, dans des bars. . . . On visite aussi des musées. Il y a un Musée d'art moderne très intéressant et gratuit pour les étudiants! Ce week-end nous allons à Paris en train. Et toi, comment ça va ? A bientôt. Bises, Audrey."

⁶ Catherine Dollez and Sylvie Pons, *Alter Ego 4 : méthode de français* (Paris: Hachette, 2007), 27: "Pluie de baisers avec battements de coeur. Persistance des frissons en fin de journée. Au cours de la nuit la fièvre atteindra la cote d'alerte et les amoureux tousseront jusqu'à l'aube."

be made up of learners of different nationalities, which can complicate interactions.

The type of cross-training activities proposed here is a response to this simultaneous opportunity and difficulty of both methods of foreign language instruction. It consists of putting learners in domestic learning situations where they are in contact with learners in immersive learning situations and getting them to work together, preferably autonomously, taking turns in the roles of trainer and learner. Its purpose is to motivate domestic learners by giving them a taste of what it is like to be immersed in a country where the language they are learning is used on a daily basis. The idea is also to motivate learners of immersive learning by making them aware of the advantages and obstacles of learning a language far from home and among speakers who already speak it.

To organize this activity, it is first necessary to establish contact with another language center. For example, *École de Guerre* in Paris has contacts with military language centers in Estonia, the United Kingdom, Austria, and even Australia.⁷ The initiative for this particular project came from an educational exchange with Estonia.

Contact should be made in advance, guided by the trainers, so that learners from both language centers understand the objectives and know how to prepare for the joint session. This is an important first contact between the two populations of learners, so that by the time the activity takes place, introductions have already been made. The activity must be carried out in a nonfacilitated videoconference, where the official trainer of each group does not intervene in conversation. They may be present, but simply as an observer and in the background. Learners need to feel fully responsible for the activity and understand this as an opportunity for them to practice speaking in the target language.

For the activity to be effective, it needs to be organized between groups of four to eight learners. Each group prepares a language-learning activity that will be proposed to the other group. Initially, the “domestic”

⁷ Estonian Military Academy, Tartu; Defense Centre for Language and Culture, Shrivenham, United Kingdom; National Defence Academy, Vienna, Austria; and Australian Defence Force School of Languages (DFSL), Canberra.

group presents its activity, conducts it with the “immersion” group, and corrects it. The roles are then reversed.

On each occasion, the group acting as “trainer” must present the objectives of the activity, lead the activity, and correct or finalize it. The activity should last between 30 and 45 minutes, which corresponds to the duration usually preferred in classes to maintain learners’ attention on a single activity, but its duration can be adapted according to the activity and learners’ level. At the end of the first activity, or in another time slot, the roles change: the learners become trainers and propose their activity. At the end, or in a third and final time slot, a discussion phase is set aside to exchange views on the activities, on their learning, and on the advantages and disadvantages of learning at home or learning in immersion. This discussion is best kept as informal as possible, and it needs no preparation: it represents an opportunity for learners to engage in the most natural and spontaneous exchange possible with their counterparts.

The proposed activities could include an oral comprehension quiz based on an authentic video, or the drafting of a deliverable based on written documents. In the first case, the learner group has no prior knowledge of the video. The group discovers it during the activity with the questionnaire. The number of viewings and the time given to answer the questions are decided by the trainer group. In the second case, the written documents are sent to the learner group beforehand. The trainer group defines the objectives of the deliverable (recipients, size, and formatting).

One previous example of a video document was an excerpt from the TV series *Karambolage* (Arte) entitled “*Les Couleurs du drapeau français*” [The Colors of the French Flag].⁸ After viewing, questions focused on lexical identification (*arborer, susurrer*), fine comprehension (Who is behind the reform led by President Emmanuel Macron?), broad comprehension (How do you justify the change from *bleu marine* to *bleu roi* under the presidency of Valéry Giscard d’Estaing?), fine interpretation (Present and give your opinion on the reaction of the French Navy.), or broad interpretation (Has your country’s flag been modified recently? Could it be modified?).

⁸ *Karambolage*, “Les couleurs du drapeau français,” directed by Claire Doutriaux, featuring Marc Chevalier and Maija-Lene Rettig, aired 2022, on ARTE France, Ilinin.

The domestic learning trainers emphasized the quality of their language learning and developed a few difficult questions to instill a playful competitive spirit with the immersive learning learners.

As an example of a document to help organize the drafting of a deliverable, the “Responding to the Challenges of Nuclear Safety” fact sheet from the French Ministry of Ecological Transition, dated 2021, could be used as a starting point for a workshop during which a note is drafted presenting the risks of civil nuclear power to French people, as well as elements to reassure them.⁹

The group of trainers could organize a short introductory session, followed by writing subgroups (one subgroup on risk factors, one subgroup on safety and security policy), and conclude the activity by pooling subgroup work. At the end of the activity, the learners’ group could finalize the deliverable and have it validated by the trainers’ group. The immersive learning trainers could use their knowledge of French society and current affairs to ensure that the deliverable is aligned with the set objective. Subgroup work provides an opportunity for discussion and debates on the theme, to make the most of the information in the basic document.

The aim of this activity is first and foremost andragogical: it attempts to put learners in the position of learning leaders by giving them the role of trainers for the duration of a session. They will have to prepare their activity in advance, with or without the help of their usual trainer. They will have to transition from learning to designing the activity, trying to anticipate difficulties, constructing questionnaires, defining workshop themes, preparing correction methods, and so on. And they will not be doing it among themselves, but with a real group of learners they do not yet know. By anticipating the difficulties, the possible right and wrong answers, and the time needed for reflection, learners put their learning into perspective so they can approach it more consciously, independently, and dynamically.

The second objective is to create a “community of learners.” The ac-

⁹ “Répondre aux défis de la sécurité nucléaire,” ICOM-DICOM/CRE/18097-1, Ministère de la Transition écologique, Service du haut fonctionnaire de défense et de sécurité, April 2021.

Table 1. An extract from “Responding to the Challenges of Nuclear Safety”

UNE APPROCHE INTÉGRÉE . . . Le service du haut fonctionnaire de défense et de sécurité se situe au carrefour de domaines régalien (contre-terrorisme, renseignement, protection du secret de la défense nationale, défense) et de domaines techniques (ingénierie nucléaire, techniques et technologies de sécurité). Il est à la confluence de l'ensemble des acteurs privés et publics liés à la sécurité nucléaire : responsables d'activité, organismes d'expertise technique et de surveillance, autres acteurs ministériels (ministères de l'Intérieur, de l'Europe et des Affaires étrangères, des Armées, Secrétariat général de la défense et de la sécurité nationale).	AN INTEGRATED APPROACH . . . The department of the senior defense and security official is at the crossroads of regal domains (counterterrorism, intelligence, protection of national defense secrets, defense) and technical domains (nuclear engineering, security techniques and technologies). It is at the confluence of all private and public players involved in nuclear security: business managers, technical expertise and monitoring organizations, other ministerial players (Ministries of the Interior, Europe and Foreign Affairs, the Armed Forces, General Secretariat for Defense and National Security).
. . . ET GLOBALE La sécurité nucléaire nécessite la prise en compte de menaces diverses et évolutives. En réponse, 3 leviers indissociables sont mis en place systématiquement : L'ANTICIPATION (réglementation, planification, intelligence économique, renseignement) ; LA PROTECTION (surveillance, dimensionnement des ouvrages de protection) ; LA DÉFINITION DU DISPOSITIF D'INTERVENTION avec les responsables d'activité et les forces de l'État et la gestion de crise. Les mesures de protection reposent sur un ensemble global de mesures passives et actives, physiques (vidéoprotection, systèmes d'accès...) ou informationnelles (protection de l'information, cybersécurité), techniques, humaines et organisationnelles. Le dispositif de sécurité est conçu en fonction des spécificités de chaque site et d'études de sensibilité et de vulnérabilité aux différentes menaces, dans un processus d'amélioration permanente. La protection des personnes et de l'environnement repose en effet sur la mise en œuvre, dans des délais satisfaisants, des actions correctives aux constats réalisés pendant les inspections ou à l'occasion de nouvelles dispositions rendues nécessaires face à l'évolution des risques identifiés. Le dispositif de sécurité repose donc sur une attitude proactive et déterminée des acteurs concernés à laquelle le service veille pour assurer l'efficacité de la politique de sécurité.	. . . AND GLOBAL Nuclear safety requires us to take into account a wide range of evolving threats. In response, three inseparable levers are systematically implemented: ANTICIPATION (regulations, planning, economic intelligence, intelligence); PROTECTION (surveillance, sizing of protective structures); DEFINITION OF INTERVENTION ARRANGEMENTS with activity managers and government forces and crisis management. Protective measures are based on a comprehensive set of passive and active physical (video protection, access systems, etc.) or informational (information protection, cybersecurity), technical, human and organizational measures. Security measures are designed to meet the specific needs of each site, and are based on studies of sensitivity and vulnerability to various threats, as part of a continuous improvement process. The protection of people and the environment depends on the timely implementation of corrective actions in response to findings made during inspections, or to new measures made necessary by changes in identified risks. The safety system therefore relies on a proactive and determined attitude on the part of the players concerned, which the department ensures to ensure the effectiveness of the safety policy.

Source: “Répondre aux défis de la sécurité nucléaire,” ICOM-DICOM/CRE/18097-1, *Ministère de la Transition écologique, Service du haut fonctionnaire de défense et de sécurité*, April 2021.

tivity at the beginning is a pretext for creating a bond of familiarity and cooperation between the two groups of learners, each of which can bring something to the other. The exchange that takes place in the last part of the activity, or in the last slot, is essential. It gives the two groups a chance to compare their difficulties, as well as the solutions they feel are best for tackling the obstacles of learning a foreign language.

This final exchange can be presented as a brainstorming session based on three fundamental questions:

- What do you find most difficult about learning French (e.g., the start of learning/during learning; to maintain one's level/to keep progressing)?
- Which activities do you find the most useful for overcoming these difficulties? Answers can be refined according to oral and written skills, and according to comprehension or production situations.
- What difficulties do you imagine, or have you encountered, when it comes to speaking French in a real situation with French speakers? While the immersive learning group is well placed to answer this question, it should also be able to correct any misrepresentations that domestic learners may have about the reality of language practice and present them with the most important sociocultural elements they need to know to be successful in speaking or writing in everyday or professional life in the target language country.

The experience at the French Language Department of the *École de Guerre* has shown that two of the main obstacles in the learning process are feelings of loneliness and uselessness. Creating a community of learners can be a very effective way for a group of learners to share their experiences with another group of learners and put into perspective the reasons for learning, the obstacles it presents, and the solutions that can make it more effective.

English-Speaking Communities of Learners: A Case Study for Transatlantic Dialogue

In the framework of the English studies curriculum at the French War College, a community of learners was launched to trigger a clash of skills for the emergence of increased interoperability. Over the years, a noticeable mismatch has evolved between the officers and the younger students. The former have a wealth of both theoretical and empirical knowledge and no mastery of debate techniques, whereas the latter are expert debaters but lack the wise analytical skills supported by life experience and many years dedicated to acquiring academic knowledge. When this asymmetry is not taken into account, the two groups can clash and come into conflict, or a so-called “a clash of skills.”

An important element of context is that the French War College, especially for its programs in English, uses a teaching approach inspired by the transformative learning theories of Jack Mezirow.¹⁰ Broadly speaking, transformative learning encourages learners to expand their consciousness to be able to question their own feelings, reactions, beliefs, assumptions, and perspectives on their purpose. The trigger for transformative learning is a disorienting dilemma (i.e., the learner is presented with a meaning structure that no longer matches their previous experience).

Learning through transformative approaches involves a change in meaning structures through reflection about content, processes, or premises. It is the transition from one meaning structure to another that allows the metanoic process to happen and the transformation to operate: the learners realize that it is not about becoming someone different but transforming their frame of reference to be able to read and take in another culture. This process fosters their ability to adapt or question their frame of reference, allowing them to critically think about their organization, nurture perspective transformation, and bring in disruptive ideas.¹¹

¹⁰ Jack Mezirow et al., *Fostering Critical Thinking in Adulthood: A Guide to Transformative and Emancipatory Learning* (San Francisco, CA: Jossey-Bass, 1991); and Jack Mezirow, *Transformative Dimensions of Adult Learning* (San Francisco, CA: Jossey-Bass, 1991).

¹¹ Jack Mezirow et al., *Transformative Learning in Practice: Insights from Community, Workplace and Higher Education* (San Francisco, CA: Jossey-Bass, 2009).

Transformative pedagogy allows the students to realize that effective learning is not about completely discarding everything that has built them until now but accepting new knowledge into their meaning structures. It is about enriching their frame of reference by bringing in new information to make it more effective in a new environment.¹²

This teaching approach sustains the possibility of critically assessing and understanding any context due to its foundational principles (i.e., critical reflection, rational discourse, centrality of experience) and allows for effective andragogical scaffolding. One of the flagship programs delivered in English at *École de Guerre* is the debate and public-speaking program.

Why Debate in a War College?

Debating was originally designed as a decision-making tool, enabling one to organize the sea of information the world has to offer to create a point of conflict from which to make an informed decision, which is at the core of a military leader's role.

Debating serves as a methodology to integrate military design and planning because it provides a setting with a high level of uncertainty and ongoing evolution. Debating intrinsically embraces notions of subjectivity and multidisciplinary. Great debaters recognize that there are diverse approaches and philosophies of debate that force us to rethink and deconstruct the core tenets of argumentation and reasoning. Debate is conceptualized as a strategic space where debaters have control over the end goal rather than a tactical space where the field is predefined. As a result, standardized methods of argumentation are inefficient in a debate since it is an activity based on competing interpretations.

A debate is a space where all participants are expected to respond to each other's arguments and make informed decisions. As a result, it is a vehicle to share diverse narratives or perspectives. It is a space where questioning is rewarded and antithetical arguments are welcomed. It can be sometimes challenging to create this type of space in a very vertical

¹² Jack Mezirow et al., *Learning as Transformation: Critical Perspectives on a Theory in Progress* (San Francisco, CA: Jossey-Bass, 2000).

hierarchical structure such as military chain of command. Subsequently, debate challenges what are generally considered unquestionable or organizational tenets and principles. It is about experimenting with what a policy or plan would look like even if it seems unrealistic in the eyes of the speaker. The discussion is not oriented around whether the plan will occur but rather on whether it should occur.

It is important to note that the approach to debate described here is the one used and taught in most English-speaking countries, but not in France where confrontation of ideas is formatted very differently. French officers have never learned debating skills or delivered debates as this is practiced in the United Kingdom or the United States. When they start the program, it is a whole new world to them, and to actually develop the skill set that goes with debate, they have to go far beyond learning the theory of debate in class with their coach. They need to engage in actual debate with those who have mastered that art of argumentation. These people can be found in British or American universities where competitive debate is practiced at the highest level.

There was a targeted outreach effort aimed at organizing debate events with U.S. universities and colleges that maintain active debate groups and traditions. The officers from *École de Guerre* were trained to prepare for and deliver within the framework of formatted debate, mainly on topics related to American policy or American parliamentary debate. They participated in events on the American university debate circuit during the academic year while on their one-year course in *École de Guerre*. This meant engaging in events organized for college students and not for senior officers attending a war college course.

The result was not very productive, as there was too much discrepancy between the two different groups. On the one hand, there were 20-year-olds who had absolutely mastered the art of debate but lacked life experience and the knowledge and wisdom that comes with it. On the other, there were majors and lieutenant colonels with exceptional life experience and achievements from which they had built a wealth of knowledge and wisdom.

This led to a very adversarial frustration, with condescending and dismissive attitudes from both sides. The young students would actively

seek to destroy the officers' cases due to the numerous gaps in their lines of argumentation, as the latter were rookies in the sport of debate and cross-examination. The officers would often then retaliate by presenting facts and evidence that highlighted the students' lack of experience and knowledge.

Despite this counterproductive situation, it was possible to find a way to move from this stalemate to an effective and creative community of learners. How were these two very different groups structured into an effective community of learners?

It was first about coming to the realization that this international exchange that had brought these two groups together was focusing on the wrong objective. The focus had been on the competitive part of debating, which entails an adversarial, and thereby potentially negative, clash. Instead of competition, it became obvious that it would be more productive to bring these two groups together in a training framework in which each side would feel comfortable showcasing and workshopping their knowledge and skills.

For this to happen, the starting point, which had the potential to become a source of friction, had to be that the American students acknowledged that the French officers held experience they did not, while the French officers had to acknowledge that these young American students had mastered a craft that they had yet to learn. The goal was not to avoid the clash but to render it productive by creating a situation in which they could teach each other the skills each had already mastered.

The goal was still to put the learners in an instructional role based on their field of expertise. The American students coached the French officers on key debating techniques such as the anatomy of an argument, warrant-building, rebuttal, analysis of competing interpretations, framing, etc., thereby increasing the officers' level of interoperability with their American counterparts for their future assignments. The French officers mentored the American students in a policy competition, using their 15+ years of experience in the field to guide them in the process of building their policy proposal. With this approach, each learner saw themselves and other learners as having the potential to be knowledge brokers and instructors. It built their awareness of existing gaps and helped them find

solutions to fill those gaps. It also built their awareness of their own qualities and their ability to share them.

This process enabled the growth of a very healthy motivation, which can be summarized in the following words: “I will become good at your craft to thank you for teaching it to me.”

To allow this to happen, an initiative was launched for an international debate module with 12 American universities. The module consisted of three events, the first one taking place in Los Angeles, California; the second in Washington, DC; and the final one in Paris. The module was named the Transatlantic Dialogues, as it studies questions that are key to both countries in matters of security and defense.¹³ The American students were competing to build the best possible case to support policies that further improved transatlantic security. They were mentored by the French officers, who used their experience to deliver useful guidance to the students on how to build policy proposals that could easily and effectively be applied to the real world. The officers then received feedback on how that guidance could be translated into a debate structure, further honing their debating skills. This project created a community that shared cultural insights, thereby bridging the key principles of andragogy and transformative learning.

Promoting the creation of communities of learners as a necessary and effective addition to andragogy can add a positive, and therefore more effective, learning experience to any course or curriculum. These communities of learners must be original and creative while relying on the learners’ initiative, independence, and autonomy as they gradually take on the roles of designers and instructors. The first approach described in this article is rooted in language learning and the other is rooted in developing leadership skills through debate in a second language. This chapter’s goal is to demonstrate that the principles of creative communities of learners will have added value in all curricula and training programs in adult learning within professional military education.

¹³ The first iteration occurred during the 2022–23 academic year.

Part 2

TRAINING & EDUCATION CASE STUDIES



CHAPTER 4 ●

What Does English Have to Do with War?

Decreasing Writing Error Propensity
in a Military Service Academy First-Year Composition Course

Lieutenant Commander Philip Garrow (USN), EdD

Ever since Ernest L. Boyer coined the phrase Scholarship of Teaching and Learning (SOTL) in 1990, it captivated educational researchers and instructors alike; no small feat considering that his initial work did not even provide a clear and concise definition.¹ About a decade and a half later, Kathleen McKinney filled that gap by conceiving of SOTL as research that “involves the systemic study of teaching and/or learning and the public sharing and review of such work through presentations or publications. Clearly, there is a focus here on the product or outcome.”² Thanks to that conclusion, SOTL was ready made for military education. A theoretical construct with a clear focus on results is a natural fit for the Armed Services predisposed to do the same thanks to the binary nature of mission success.

Service academy populations present a unique research opportunity given that they are what the late Jana L. Pershing called “closed sociological systems,” or unique groupings of like-minded citizens surrounded by

¹ Ernest L. Boyer, *Scholarship Reconsidered: Priorities of the Professoriate* (Princeton, NJ: Carnegie Foundation for the Advancement of Teaching, 1990).

² Kathleen McKinney, “The Scholarship of Teaching and Learning: Past Lessons, Current Challenges, and Future Visions,” *To Improve the Academy* 22, no. 1, (June 2004): 8, <https://doi.org/10.1002/j.2334-4822.2004.tb00399.x>.

fences and armed guards.³ Yet, this near-captive audience also presents a special challenge for institutional review boards. How does an officer who is superior in rank to those they instruct ever conduct research on their own students without inadvertently pressuring them to participate? SOTL methods are one means through which one can balance the requirements of research with the demands of the classroom.⁴

From 2020 to 2021, a mixed methods SOTL study was conducted at the U.S. Naval Academy (USNA) to analyze professional communication instruction in a first-year composition (FYC) course. The study occurred in two phases, with 55 midshipmen participating in a needs assessment data gathering initiative to evaluate their volume of formal writing (i.e., grammar or spelling) and citation errors. To simulate the different professional writing style guides used in various U.S. Navy fleets, multiple documentation styles—Modern Language Association (MLA), American Psychological Association (APA), and the *Chicago Manual of Style* (CMS)—were required for the FYC’s writing assignments. A total of 55 students allowed their coursework to be examined; they demonstrated an average errors per 100 words rate of .29 for formal problems as well as .65 for documentation style issues.⁵ During the intervention stage, the FYC was changed to include scaffolded writing assignments, in-class reflective activities, and the inclusion of a military-themed novel. As a result of those course design alterations, another set of 50 midshipmen participants lowered their errors per 100 rates for both citation issues (.77 to .62) and formal writing errors (.58 to .23).

Background

Military missions inherently involve elevated levels of risk, often exacerbated by specific environments; these challenges require a talented officer

³ Jana J. Pershing, “Whom to Betray?: Self-regulation of Occupational Misconduct at the U.S. Naval Academy,” *Deviant Behavior* 23, no. 2 (2002): 155, <https://doi.org/10.1080/016396202753424538>.

⁴ Andrea S. Webb and Ashley J. Welsh, “Phenomenology as a Methodology for Scholarship of Teaching and Learning Research,” *Teaching & Learning Inquiry* 7, no. 1 (2019): 168–81, <https://doi.org/10.20343/teachlearninqu.7.1.11>.

⁵ *Midshipmen* refers to a gender-neutral military rank. Meredith Newman, “Navy: ‘Midshipman’ Title Will Not Change,” *Capital Gazette*, 9 December 2016.

corps familiar with critical thinking and clear communication practices.⁶ The danger associated with tiny professional writing errors is great, up to and including unnecessary deaths.⁷ All five U.S. military Service academies (the single biggest commissioning source for each branch) make the FYC a requirement for graduation.⁸ That undergraduate students make writing mistakes is not late-breaking news; for more than 100 years, scholars have analyzed college papers to see where learners seem to stumble the most.⁹ Similarly (but less commonly), research also shows that undergraduates find citation standards difficult to follow.¹⁰

The challenge for a FYC instructor at a Service academy is convincing students that such commonplace problems are much more dire when they occur within their perfection-conscious chosen profession. To this end, treating them as adults, possessed of agency and awareness, is a useful tactic. A practical aspect of Malcolm Knowles' characterization of

⁶ Fred W. Kacher, *Newly Commissioned Naval Officer's Guide*, 2d ed. (Annapolis, MD: Naval Institute Press, 2018); Lisa A. McComas and Joshua D. Kristenson, *The Naval Officer's Guide*, 13th ed. (Annapolis, MD: Naval Institute Press, 2019); and J. Michael Rifenburg, *Drilled to Write: Becoming a Cadet Writer at a Senior Military College* (Louisville, CO: Utah State University Press, an imprint of University Press of Colorado, 2022), 19.

⁷ J. Michael Rifenburg, "To Ensure Warfighting Function: Writing Inside a U.S. Army Brigade Headquarters," *Composition Studies* 47, no. 1 (2019): 116–34; J. Michael Rifenburg and Brian G. Forester, "First-Year Cadets' Conceptions of General Education Writing at a Senior Military College," *Teaching & Learning Inquiry* 6, no. 1 (2018): 52–66, <https://doi.org/10.20343/teachlearningqu.6.1.6>; and Robert Shenk, *The Naval Institute Guide to Naval Writing* (Annapolis, MD: Naval Institute Press, 2008).

⁸ Kristen Moon, "What Students Need to Know Before Applying to One of the United States Military Academies," *Forbes*, 20 August 2019.

⁹ Roy I. Johnson, "The Persistency of Error in English Composition," *School Review* 25, no. 8 (October 1917): 555–80, <https://doi.org/10.1086/436796>; Paul A. Witty and Roberta La Brant Green, "Composition Errors of College Students," *English Journal* 19, no. 5 (May 1930): 388–93, <https://doi.org/10.2307/803405>; and Robert J. Connors and Andrea A. Lunsford, "Frequency of Formal Errors in Current College Writing, or Ma and Pa Kettle Do Research," *College Composition and Communication* 39, no. 4 (1988): 395–409, <https://doi.org/10.2307/357695>.

¹⁰ Aaron S. Blicblau, Megan Bruwer, and Kevin Dini, "Do Engineering Students Perceive that Different Learning and Teaching Modes Improve Their Referencing and Citation Skills?," *International Journal of Mechanical Engineering Education* 44, no. 1 (2015): 3–15, <https://doi.org/10.1177/0306419015624186>; Marni Davis and Jill E. Anderson, "Follow the Footnote," *History Teacher* 52, no. 3 (May 2019): 477–98; and Xiaodong Zhang, "Teaching Citations/Referencing: How Do Chinese College Student Writers Respond?," *Publishing Research Quarterly* 34, no. 4 (2018): 580–94, <https://doi.org/10.1007/s12109-018-9615-y>.

andragogy, a formal term for adult learning, is that older students find it helpful to understand how or why a given piece of information may relate to their profession.¹¹ Using real-world military writing examples and exploring subjects such as strategy, tactics, and discipline in an FYC (e.g., incorporating in-class reflective writing activities) allows officer candidates the chance to draw connections between coursework and the battlefield.¹² SOTL is at its finest when it facilitates instructors and students exploring together the nature and practicality of learning.

Literature Review

A literature review was compiled before commencing data collection. This step was especially critical given that, as Alicia Capello and Janice Miller-Young pointed out, “SOTL has been criticized for weak literature reviews.”¹³ Some of the individual factors explored included: overall learner attentiveness; the ability of students to comprehend what they read; the importance of following existing standards and widely recognized rules in professional writing; and, most importantly, the causes and repercussions—both real and potential—of human error.¹⁴ Industries examined in the literature review expanded beyond military affairs by also looking for causes and considerations in writing errors afflicting commercial busi-

¹¹ Malcolm Knowles, *The Modern Practice of Adult Education: From Pedagogy to Andragogy*, 2d ed. (New York: Cambridge Books, 1980); and Malcolm Knowles and Associates, *Andragogy in Action: Applying Modern Principles of Adult Learning* (San Francisco, CA: Jossey-Bass, 1984).

¹² Rifenburg and Forester, “First-Year Cadets’ Conceptions,” 53.

¹³ Alicia Capello and Janice Miller-Young, “Who Are We Citing and How?: A SOTL Citation Analysis?,” *Teaching & Learning Inquiry* 8, no. 2 (2020): 2, <http://dx.doi.org/10.20343/teachlearninqu.8.2.2>.

¹⁴ Katie Greer and Sean McCann, “Everything Online Is a Website: Information Format Confusion in Student Citation Behaviors,” *Communications in Information Literacy* 12, no. 2 (November 2018): 150–65; Adrienne E. Barnes and Young-Suk Kim, “Low-skilled Adult Readers Look Like Typically Developing Child Readers: A Comparison of Reading Skills and Eye Movement Behavior,” *Reading and Writing: An Interdisciplinary Journal* 29, no. 9 (May 2016): 1889–914, <https://doi.org/10.1007/s11145-016-9657-5>; Timothy S. Lesar, Laurie Brice-land, and Daniel Stein, “Factors Related to Errors in Medication Prescribing,” *JAMA* 277, no. 4 (1997): 312–17, <https://doi.org/10.1001/jama.1997.03540280050033>; and J. Bryan Sexton, Eric J. Thomas, and Robert L. Helmreich, “Error, Stress, and Teamwork in Medicine and Aviation: Cross Sectional Surveys,” *Journal of Human Performance in Extreme Environments* 6, no. 1 (2001): 6–11, <https://doi.org/10.7771/2327-2937.1019>.

ness, medical practice, and specialty fields such as commercial shipping and aviation.

Such disparate areas were tied together by employing Jennifer Watling Neal and Zachary P. Neal's 2013 update to Urie Bronfenbrenner's ecological systems theory.¹⁵ Rather than rely on the classic model of envisioning human relationships as a series of concentric circles in which various systems, such as the microsystem or the exosystem, are contained within one another, the Neals creatively argue that what they term a *networked* model is superior in the information age.¹⁶ By separating the systems yet showing how they are still tied together by individual relationships (and therefore can and sometimes do still overlap, as in the nested model), the Neals' research provides a more accurate representation of the possible effects on a singular learner in the digital age. While the Neals do not mention SOTL by name, their innovative approach to rethinking a major foundational theory is certainly within the SOTL spirit.

Needs Assessment Methodology

After obtaining Institutional Review Board approval for human research from both USNA as well as Johns Hopkins University (the study was undertaken as part of a doctoral dissertation), a needs assessment was completed to assess the nature of midshipmen's formal and citation error patterns. The FYC at USNA is divided into fall and spring semesters, with the great majority of students usually assigned a different instructor (sometimes military, sometimes civilian) for each half.¹⁷ The researcher (a military instructor) employed convenience sampling and sought participants from students already enrolled in the researcher's classes.¹⁸ Although the original intention was to only solicit participants from the

¹⁵ Urie Bronfenbrenner, *The Ecology of Human Development: Experiments by Nature and Design* (Cambridge, MA: Harvard University Press, 1979).

¹⁶ Jennifer Watling Neal and Zachary P. Neal, "Nested or Networked?: Future Directions for Ecological Systems Theory," *Social Development* 22, no. 4 (November 2013): 728, <https://doi.org/10.1111/sode.12018>.

¹⁷ "Periodic Program Review" (unpublished manuscript, Department of English, U.S. Naval Academy, March 2020), 5–6.

¹⁸ Chad R. Lochmiller and Jessica N. Lester, *An Introduction to Educational Research: Connecting Methods to Practice* (Thousand Oaks, CA: Sage, 2017), 141.

spring 2020 semester, the COVID-19 pandemic necessitated a transition to remote learning midsemester and severely impacted the number of midshipmen who elected to participate (12 out of 35, or 34.2 percent). As a result, the needs assessment was replicated when classes returned to in-person in the fall of 2020, with many more volunteers (43 out of 54 invitees, or 79.6 percent) submitting useable materials for analysis.

Although demographic data was not a key element of the needs assessment information collected, it is worth mentioning that the composition of the FYC classes used during this period (as well as the graduating class that later participated in the intervention) roughly approximated the statistics of their graduating classes as a whole. In general, about 70 percent of each class was male compared to 30 percent female (intersex data was unavailable). Meanwhile, roughly 60–70 percent of the various graduating classes identified as White, contrasted with 12–15 percent who selected Hispanic, the 8–9 percent that listed Asian American, or the 7 percent of the group that choose African American, as well as students who identified with multiple races.¹⁹ No attempt was made to analyze or correlate error propensity with a given participants gender, race, or any other demographic consideration (i.e., whether the student had previously served in a branch of the armed forces or whether they were the child of alumni).

Of note, a complication of SOTL in a military environment is that students working with a military FYC instructor might feel unnecessarily compelled to participate in a voluntary study (perhaps out of a sense of obligation to the superior rank of the instructor or the institutional emphasis on being a team player). To reduce that possibility, the researcher asked one of the students to distribute and collect consent forms without the instructor present and to later provide them to the researcher's supervisor for safekeeping until the end of the semester. Only after grades were finalized was the researcher allowed to commence data analysis. An additional benefit was that the in-class learning experience was identical for participants and nonparticipants alike.

¹⁹ Philip Garrow, "Creativity and Conformity in First-Year Composition at the U.S. Naval Academy" (EdD diss., Johns Hopkins University, 2022), 44.

Information was obtained during the needs assessment stage by tabulating various types of errors on assignments that student participants submitted for the class as well as by analyzing and coding ungraded (or low-stakes) in-class writing activities. Some key concepts that relate to professional writing emerged during this process and informed the language of the research questions, to include: andragogical principles; student “learning-orientation” of pupils; universal design of the course curriculum design; and reflective writing.²⁰ Four needs assessment questions organized the research along both quantitative and qualitative lines to adopt a mixed methods approach.

RQ1: How do the respective averages of grammar and citation “errors per 100 words” on FYC course assignments submitted by midshipmen compare to earlier those featured in earlier studies of undergraduate writing?

RQ2: Do differences in the respective averages of grammar and citation “errors per 100 words” on FYC course assignments submitted by midshipmen correlate to their self-identification as either high or low learning- and/or grade-oriented students?

RQ3: How are midshipmen assessments of meaningfulness, competence, impact, and choice reflective of their sense of learner empowerment in the FYC classroom?

RQ4: What terms and concepts do midshipmen use to characterize their experiences and expectations re-

²⁰ Thomas Kent, “The ‘Remapping’ of Professional Writing,” *Journal of Business & Technical Communication* 21, no. 1 (2007): 12–14, <https://doi.org/10.1177/1050651906293506>; Knowles and Associates, *Andragogy in Action*; James A. Eison, “A New Instrument for Assessing Students’ Orientations Towards Grades and Learning,” *Psychological Reports* 48, no. 3 (1981): 919–20, <https://doi.org/10.2466/pr0.1981.48.3.91>; James A. Eison, Howard R. Pollio, and Ohmer Milton, “Educational and Personal Characteristics of Four Different Types of Learning- and Grade-oriented Students,” *Contemporary Educational Psychology* 11, no. 1 (1986), 55, [https://doi.org/10.1016/0361-476X\(86\)90012-3](https://doi.org/10.1016/0361-476X(86)90012-3); Robert Rotenberg, *The Art & Craft of College Teaching*, 2d ed. (New York: Routledge, 2016), 86; and Linda Nilson, *Teaching at Its Best: A Research-based Resource for College Instructors*, 4th ed. (San Francisco, CA: Jossey-Bass, 2016), 233.

garding the importance of the FYC curriculum in their future military career?²¹

Error data was gleaned from several specific course assignments, notably three distinct varieties of student papers: a textual analysis (as well as a contextual analysis) along with a response assignment (during which a student analyzed a classmate's analysis paper).²² Each paper required a separate citation format (MLA, APA, and CMS, respectively). Three in-class activities involving the completion of reflective surveys served as the other source of participant information. Likert-scale-only instruments used during the survey activities included the LOGO- II and the learner empowerment instrument (LEI).²³ The third instrument, the 1988 IDEA Survey, featured a section for open answers as well as Likert responses.²⁴ Information obtained from these surveys and student course work yielded five measurable constructs: formal errors, citation errors, learning orientation, learner empowerment, and career relevance.²⁵

Findings from the Needs Assessment

With respect to RQ1, needs assessment participants made fewer formal mistakes for every 100 words (.29) than the students associated with the earlier 1988 Robert J. Connors and Andrea A. Lunsford (2.26) study cit-

²¹ Garrow, "Creativity and Conformity in First-Year Composition," 41.

²² Jack Selzer, "Rhetorical Analysis: Understanding How Texts Persuade Readers," in Charles Bazerman and Paul Prior, eds., *What Writing Does and How It Does It* (Mahwah, NJ: Earlbaum, 2004), 283.

²³ The original survey instrument from 1981 was called LOGO, which stood for Learning-Oriented Grade-Oriented. But when they published the 1983 manual for LOGO-II, no such definition for the acronym was provided. James A. Eison, Howard R. Pollio, and Ohmer Milton, "Manual for Use with LOGO-II," in *Theory and Practice in Teacher Education Publications and Other Works* (Knoxville: University of Tennessee, 1983), 6–8; and Andrea Bainbridge Frymier, Gary M. Shulman, and Marian Houser, "The Development of a Learner Empowerment Measure," *Communication Education* 45 no. 3 (July 1996): 187, <https://doi.org/10.1080/03634529609379048>.

²⁴ Donald P. Hoyt and Eun-Joo Lee, *Basic Data for the Revised IDEA System*, IDEA Technical Report 12 (Manhattan: Kansas State University, 2002), 62–63.

²⁵ Marian Houser and Andrea Bainbridge Frymier, "The Role of Student Characteristics and Teacher Behaviors in Students' Learner Empowerment," *Communication Education* 58, no. 1 (January 2009): 35–53, <https://doi.org/10.1080/03634520802237383>.

ed earlier or the 2008 replication of the 1988 research performed by Andrea Lunsford and Karen J. Lunsford (that investigation featured an error rate of 2.45).²⁶ Yet, the midshipmen's needs assessment essays featured 4.09 errors per paper on average, a figure far too high for the zero-defect mentality of the military. The volume of documentation style mistakes was more severe; error-per-100-words rates across the assignments ranged from .49 to .95 with 7.6–9.31 average total errors per paper. Given that midshipmen's citation problems were roughly twice the formal total, the implications for their ability to compose error-free professional documents appeared even graver. For RQ2, no distinct relationship between midshipmen's formal or citation mistakes and their learning orientation (in any of the four categories identified in earlier studies) emerged; therefore, the takeaway was that any effective intervention should apply equally to students regardless of where they fell on the learning orientation spectrum.²⁷

Results associated with RQ3 showed more promise. By compiling descriptive statistics from LEI survey responses, it was readily apparent that the means for some of the five-point Likert scale items—notably *meaningfulness* (2.75) and *impact* (2.78)—were more important to midshipmen than constructs like *competence* or *choice* (by a difference of about .5). In-vivo coding was applied to the free writing sections of the IDEA survey to respond to RQ4; the key concepts that emerged suggested that any course adjustments should involve *professional development*, an emphasis on *real-life situations*, and the inclusion of *more feedback*.

Results from all four research questions were used to design a *cognitive-apprenticeship* style intervention, a method that favors a combination of activity and thoughtful reflection to maximize the odds of long-term memory retention with respect to learning.²⁸ In many ways,

²⁶ Andrea A. Lunsford and Karen J. Lunsford, “‘Mistakes Are a Fact of Life’: A National Comparative Study,” *College Composition and Communication* 59, no. 4 (June 2008): 800.

²⁷ Eison, Pollio, and Milton, “Educational and Personal Characteristics,” 55.

²⁸ John Seely Brown, Allan Collins, and Paul Duguid, “Situated Cognition and the Culture of Learning,” *Educational Researcher* 18, no. 1 (1989): 32–42, <https://doi.org/10.3102/0013189x018001032>; and Peter C. Brown, Henry L. Roediger III, and Mark A. McDaniel, *Make It Stick: The Science of Successful Learning* (Cambridge, MA: Harvard University Press, 2014).

this approach mirrors the hands-on learning style commonly employed in military commands.²⁹ By asking students to complete writing activities while considering how and why such knowledge might be useful to them in the future, instructors can benefit from the positive aspects of varied practice.³⁰ For example, employing different documentation styles in the classroom helps prevent a default to rote memorization. Cognitive apprenticeship holds much in common with critical thinking in that both ask learners to not just absorb but rather reflect and engage with what knowledge acquisition means in their lives and careers.

Intervention Methodology

Given how few USNA students are assigned the same FYC instructor for both semesters, a longitudinal study was not possible. For the intervention stage, 50 new volunteers (a 70.4 percent response rate from the available group) were obtained using the same IRB-sanctioned coercion safeguards used during the needs assessment. As before, these members of the class of 2025 approximated their graduating class as a whole sufficiently so as to marginalize demographic considerations. Also as earlier, SOTL served as the intervention stage conceptual framework because it accounted for both the role of the instructor and that of the student in an intellectual growth environment.³¹ But a number of other concepts (as defined by earlier researchers) were considered and incorporated as subelements in the effort to reduce student citation and formal error propensity. Said concepts included: Robert H. Ennis' definition for critical thinking; John Dewey's description of inquiry-based learning; Ellen Cushman's exploration of the role of rhetoric instructors with respect to social justice; Linda Nilson's approach to reflective writing; Linda Bradley's standards for peer review; and Zane K. Quible's focus on the importance of error

²⁹ Kacher, *Newly Commissioned Naval Officer's Guide*.

³⁰ Brown, Roediger, and McDaniel, *Make It Stick*, 47.

³¹ Pat Hutchings, "Theory: The Elephant in the Scholarship of Teaching and Learning Room," *International Journal for the Scholarship of Teaching and Learning* 1, no. 1 (2007): 2, <https://doi.org/10.20429/ijSOTL.2007.010102>.

labeling.³² These concepts all informed the cognitive-apprenticeship approach to the intervention because they were explained to the participants as related elements of the andragogical approach to the study. Some research questions employed during the intervention stage closely mirrored those of the needs assessment, but three more were added to help assess the effectiveness of using SOTL in a mixed methods study.

RQ1: To what extent, if any, did the formal grammar and mechanic “error rate per 100 words” change for midshipmen during the course of the semester?

RQ2: To what extent, if any, did the citation “error rate per 100 words” change for midshipmen during the course of the semester?

RQ3: Do differences in the respective averages of grammar and citation “errors per 100 words” on FYC course assignments submitted by midshipmen correlate to their self-identification as either high or low learning- and/or grade-oriented students?

RQ4: What difference, if any, exists between the way(s) participants respond to the intervention assignment and class activity changes?

RQ5: How do participants describe their learning experience as well as their perception of the instructor’s role?

RQ6: How did participant sense of learner empowerment change after participating in the intervention?

RQ7: How, if at all, did midshipmen’s perception

³² Robert H. Ennis, “Critical Thinking Assessment,” *Theory into Practice* 32, no. 3(1993): 179–86, <https://doi.org/10.1080/00405849309543594>; John Dewey, *Experience and Education* (Indianapolis, IN: Kappa Delta Pi, 1938; New York, NY: Free Press, 1997); Ellen Cushman, “The Rhetorician as an Agent of Social Change,” *College Composition and Communication* 47, no. 1 (1996): 7–28, <https://doi.org/10.2307/358271>; Nilson, *Teaching at Its Best*, 233; Linda Bradley, “Peer-reviewing in an Intercultural Wiki Environment—Student Interaction and Reflections,” *Computers and Composition* 34 (2014): 80–95, <https://doi.org/10.1016/j.compcom.2014.09.008>; and Zane K. Quible, “Impact of Error Labeling on Error Elimination in Business Writing,” *Business Communication Quarterly* 69, no. 1 (2006): 8–24, <https://doi.org/10.1177/1080569905285597>.

of writing in the military change after revising the three FYC assignments?³³

The intervention model selected for this stage was convergent design to allow for data analysis by collection set (needs assessment data versus intervention) before evaluating the impacts of intervention changes on student error rates.³⁴ Process evaluation was also considered while analyzing the data sets: the midshipmen's responsiveness was analyzed in three questions (RQ1, RQ2, and RQ4), while the program delivery's quality was considered in RQ5 and the military-specific context of the investigation was factored into RQ7.

Three major changes to the FYC course design—informed by the needs assessment findings—served as intervention steps. First, to capitalize on andragogy's focus on career relevance, a required reading assignment centered on military operations and veterans' mental health issues was used in place of a different volume employed in the needs assessment that did not touch on those things. Second, additional time was allotted in class for midshipmen to engage in reflecting writing. Third, what were once individual assignments without any instructor follow-up were changed to require the submission of a first and a revised draft. This scaffolded approach allowed for a larger data set and facilitated a "one-group pretest-posttest design using a double pretest."³⁵ Such a design enabled the comparison of error propensity over time as students completed various written projects throughout the length of the course.

Beyond the midshipmen participants' writing assignments and reflections, the same needs assessment phase instruments—the LOGO-II, the LEI, and the 2002 IDEA survey—were again used to collect both qualitative and quantitative data sources. Participant responsiveness was determined via the percentage of students that elected to complete the

³³ Garrow, "Creativity and Conformity in First-Year Composition," 86–87.

³⁴ John W. Creswell and Vicki L. Plano Clark, *Designing and Conducting Mixed Methods Research*, 3d ed. (Thousand Oaks, CA: Sage, 2018), 73.

³⁵ William R. Shadish, Thomas D. Cook, and Donald T. Campbell, *Experimental and Quasi-Experimental Designs for Generalized Causal Inference* (New York: Houghton Mifflin, 2002), 110.

nongraded reflective writing activities; a mere 3 out of 50 midshipmen chose to opt out. To determine delivery quality, descriptive statistics were calculated from the quantitative-answer portion of the 2002 IDEA Survey; 68 percent of Likert-scale responses to questions asking about the instructor and the course design boasted 75 percent matching medians, while 70 percent showed the same three-quarters positive answer rate with respect to mode. Those rates strongly suggest that midshipmen responded to the instructor essentially equally regardless of which class section they were assigned, an indication that the education delivery provided to each group was approximate. To gauge how well midshipmen understood the context of the course, descriptive statistics were applied to the 2002 IDEA survey Likert-scale item assessing relevance; in this area, the addition of a military-themed reading assignment had little impact, given that the mean and the mode answers to that question for both the needs assessment and intervention participants were identical.

Findings from the Intervention

How well did the intervention work at reducing student error? Regarding RQ1 (which dealt with formal errors), intervention stage participants committed more formal writing errors per 100 words (1.20) as well as a greater volume of average of total errors (5.55) than the needs assessment group (.29 and 4.09). Although distressing, the cause of that disparity did not emerge from the data. It stands to reason, however, that the massive shift to distance learning necessitated by COVID-19 made the experience of different high school graduating classes particularly distinct.³⁶ Information literacy, always difficult to impart, became even more problematic when so many students across the country were physically unable to visit their school library. Admittedly, though, the midshipmen were less error prone even prior to the intervention than their civilian counterparts who participated in previous studies over the last century (table 1).

³⁶ Sumitra Pokhrel and Roshan Chhetri, "A Literature Review on Impact of COVID-19 Pandemic on Teaching and Learning," *Higher Education for the Future* 8, no. 1(2021): 133–41, <https://doi.org/10.1177/2347631120983481>.

Table 1. Comparison of error rates per 100 words post-intervention

Study	Year	Average length	Errors per paper	Errors per 100 words
Johnson	1917	162 words	3.42	2.11
Witty and Green	1930	231 words	5.18	2.24
Connors and Lunsford	1986	422 words	9.52	2.26
Lunsford and Lunsford	2006	1,038 words	-	2.45
Garrow (needs assess)	2020	1,411 words	4.09	0.29
Garrow (intervention)	2022	1,488 words	5.55	1.20

Source: Garrow, “Creativity and Conformity in First-Year Composition,” 115.

But most promisingly, the propensity to make mistakes decreased from the beginning to the end of the FYC completed by the intervention stage graduating class, with 8.78 total mistakes on the first assignment issued at the start of the course down to 3.88 on the final essay due by the conclusion. Citation error reduction rates examined in response to RQ2 were not as impressive. As before, the intervention students committed larger volumes of documentation style mistakes on each assignment than did the needs assessment participants (COVID was likely at play in this instance, as well). But citation errors at least mirrored the formal writing mistake pattern of becoming less frequent as the course continued, from a high total average mistake rate of 11.36 down to 10.4 as well as the average-per-100-words rate shifting from .77 to .62 from the first assignment to the last.

Findings were more mixed among the latter research questions. The learning orientation considerations associated with RQ3 remained just as opaque as they did during the needs assessment; no clear pattern arose during data assessment, which may explain why so few twenty-first-century studies utilized LOGO-II as a collection instrument. In this case study, no useful correlation between learning orientation and error propensity could be drawn. The final takeaway—that a given instructor should be prepared to cater to students who care more about grades

versus those who put a premium on knowledge acquisition for its own sake—is not particularly original or unexpected.

In a more positive vein, though, the andrological implications of this SOTL work proved more promising while evaluating RQ4. There, 8.74 percent of midshipmen's open-ended comments expressed their perspective of how well their FYC was tailored to their upcoming military career. A second indication of their enthusiasm for the course design changes was that 58 percent of midshipmen chose to write an essay on the class's military novel, while only 3.63 percent of the needs assessment group wrote about the nonmilitary novel that was replaced during the intervention. RQ5 revealed a mismatch between the relatively balanced responses to the LOGO-II, where most students clustered in the middle of the self-identified learning orientation pack by only slightly valuing grades over learning or vice versa, while the same group's responses on the 2002 IDEA Survey greatly favored discussing the importance of learning (65.63 percent) over class scores (11.465 percent).

The reasons for the disparity are difficult to discern but could be that midshipmen were not in a mindset concerned with scores while they were completing a reflective writing activity that they knew was ungraded. Further, the influence of the researcher is likely to have impacted midshipmen's responses. Midshipmen were informed that the in-class survey activities were not anonymous, so that participants could eventually be separated from nonparticipants, but this lack of anonymity perhaps resulted in less candor than might otherwise have emerged. It further seems possible that Service academy students would feel a need to provide modified answers to an officer of a much higher rank given the Naval Academy's institutional emphasis placed on demonstrating respect for superiors.

Student survey responses used for the penultimate research question also proved unpromising; the mean and median answers to the LEI Likert-scale items were so similar between the needs assessment and intervention groups as to suggest that the intervention course design adjustments had no practical effect. This high degree of similarity, though, at least justified employment of convergent design, increasing validity and reliability through the use of two separate test groups. The final research

question, dealing with the Armed Services as a context, indicated that the midshipmen's appreciation of the FYC to their chosen career improved as the semester progressed. The volume of qualitatively coded comments addressing military usefulness more than doubled (6.71 percent to 14.48 percent) between the first and final reflective writing prompt replies.

Conclusions

This study found that the intervention steps incorporating scaffolding, reflective writing, and a book relevant to the participants' career did facilitate cognitive apprenticeship and increase the ability of adult learners to better appreciate the practical benefits of the FYC. Measuring and assessing student learning is an age-old problem in education research that is not exclusive to SOTL studies. Because of its multipronged approach, mixed methods research is uniquely positioned to demonstrate intellectual growth.

Although multiple research questions were explored during both the needs assessment and intervention stages, the overarching goal of those question sets was to examine—and later reduce—midshipmen's propensity to commit formal writing and citation mistakes. With respect to decreasing error propensity, the intervention steps achieved their intended goal. As shown in the following tables, the quantitative reduction in the total number of formal (8.78 to 3.88) and documentation style errors (11.36 to 10.4) during the length of the intervention-stage graduating class indicated the success of the intervention's changes to the FYC course design.³⁷

Recall that midshipmen were required to submit assignments in the order shown in the table, with the textual analysis first and the contextual piece last. Formal writing errors consistently decreased over time, although the drop in formal writing errors between the second and third assignments was comparatively slight. With respect to citation errors, the error rate from the first to last assignment did drop, allowing the intervention to be declared worthwhile. Yet, the explanation for the rise in citation error rates on the response assignment remained unclear. Notably, midshipmen seemed more adept at decreasing formal writing errors

³⁷ Garrow, "Creativity and Conformity in First-Year Composition," 116–17.

Table 2. Intervention midshipmen formal writing error rates per 100 words

Type	Style	Average length	Formal writing errors per paper	Errors per 100 words
Textual	MLA	1,575 words	8.78	.58
Response	CMS	1,023 words	4.00	.39
Contextual	APA	1,865 words	3.88	.23

Source: Garrow, “Creativity and Conformity in First-Year Composition,” 116.

Table 3. Intervention midshipmen citation error rates per 100 words

Type	Style	Average length	Citation errors per paper	Errors per 100 words
Textual	MLA	1,575 words	11.36	.77
Response	CMS	1,023 words	12.9	1.26
Contextual	APA	1,865 words	10.4	.62

Source: Garrow, “Creativity and Conformity in First-Year Composition,” 117.

as opposed to documentation style errors despite the instructor’s greater emphasis on the latter. In the researcher’s opinion, grammar and spelling tools are much more user-friendly in common text processors like Microsoft Word or Google docs than are their citation counterparts in software such as Zotero. Although students did not mention it in their reflection, it seemed reasonable to speculate that ease-of-use for word processing tools played a role in this studies’ results.

In a more positive vein, the increasing amount of focus on military applicability in the reflections indicates that participants internalized the significant dangers presented by writing errors in future military correspondence. Knowles’ core conception of andragogy as a means to motivate adult students by focusing on professional usefulness was upheld. But since the reflections and survey results were not anonymous (to help the researcher isolate submissions from students who elected not to participate), it is possible that participants chose to emphasize positive aspects of the FYC experience rather than using those tools as a means to provide constructive criticism. If this study is replicated, it would be

advantageous to ensure that a means to provide anonymous survey responses to the researchers is incorporated.

Despite the challenges outlined above, the merits of SOTL as a powerful tool for the investigation of military issues was reinforced by the study's findings. SOTL is unique in that it is a pedagogical approach that values widely recognized techniques while also placing a premium on the benefits of innovation.³⁸ Although efforts to better define SOTL are ongoing, its usefulness as a research category is recognized. Thomas G. Ryan, for example, observed how critical it is that SOTL not be arbitrarily lumped in with related concepts like action research.³⁹ Scholars like J. Michael Rifenburg have long recognized the value of SOTL as it relates to the study and analysis of members of the armed forces.⁴⁰

Future investigations would do well to consider how the intertwined roles of military superiors and subordinates sometimes mirror and amplify the natural relationship between teachers and students in ways that SOTL is uniquely poised to help illuminate. Ideally, a similar study of formal writing and documentation-style error propensity could be undertaken at the other Service academies to see if this research study's findings and conclusions are unique to midshipmen or if they are appropriate for other cadets. Samuel P. Huntington's seminal 1957 text *The Soldier and the State* was among the first to explore the individual cultures fostered by each American Service branch, and the distinctions have remained significant during the subsequent decades.⁴¹ It would be fascinating to see if significant differences in Service academy student writing error and citation rate propensity emerged, especially if the data shed light on whether

³⁸ Clarke Mathany, Katie M. Clow, and Erin D. Aspenlieder, "Exploring the Role of Scholarship of Teaching and Learning in the Context of the Professional Identities of Faculty, Graduate Students, and Staff in Higher Education," *Canadian Journal for the Scholarship of Teaching and Learning* 8, no. 3 (December 2017): 1–16, <https://doi.org/10.5206/cjsOTL-rca-cea.2017.3.10>.

³⁹ Thomas G. Ryan, "The Scholarship of Teaching and Learning within Action Research: Promise and Possibilities," *Inquiry in Education* 4, no. 2 (November 2013): 1.

⁴⁰ Rifenburg and Forrester, "First-Year Cadet's Conceptions," 52; and Rifenburg, *Drilled to Write*, 19.

⁴¹ Samuel P. Huntington, *The Soldier and the State: The Theory and Politics of Civil-Military Relations* (Cambridge, MA: Belknap Press, an imprint of Harvard University Press, 1957).

the spike in CMS errors is a persistent feature of the intervention or if that was unique to the participant group examined here.

Given that each institution touts the merits of critical thinking and FYC's natural role in developing it, a collective embrace of SOTL could prove particularly fruitful.⁴² The comparative data shown in table 1 suggests that the USNA's emphasis on precision, professionalism, and hard work contributes to their lower degree of writing error propensity when contrasted with civilian student populations. Yet, the intervention steps demonstrated that those rates can be lowered even further when an emphasis on formal writing and citation error reduction is scaffolded into a FYC class. SOTL methods do not guarantee perfection, but they remain powerful tools for instructors interested in striving for it. A joint commitment to fostering SOTL research by the various Service academy leadership teams would go a long way in advancing Jana Pershing's earlier investigations into what makes military campus cultures sociologically unique while also helping prepare future officers for risky operational environments where professional correspondence mistakes are treated with severity.⁴³

⁴² Dana H. Born, Andrew T. Phillips, and Timothy E. Trainor, "America's Service Academies, Your Service Academies," *Liberal Education* 98 no. 1 (2012): 46.

⁴³ Pershing, "Whom to Betray," 155.



CHAPTER 5 ●

Closing the Gap between Intellect and Character in an Online Ethics Course

Colonel Heather Smigowski (USA), PhD

Students learn ethical theories and practice decision making during the leadership ethics online elective course at the U.S. Army War College (USAWC). All students apply critical thinking through written, asynchronous submissions and demonstrating intellectual understanding of the principles applied to case studies; however, a disparity was noted during virtual discussions and the resident version of the course. Students consistently demonstrated that they know the right actions to take, but they admit that they would not always perform those actions.

This concern is reminiscent of Robert Coles' 1995 essay "The Disparity between Intellect and Character," which espoused a need to reimagine ethics in education.¹ This begs the question of whether anything has changed in education since there is still a disparity between knowing what is right and doing what is right when faced with a complex situation.

To close the gap between student thoughts and actions in the USAWC ethics course, reflective exercises were introduced. Initial results indicate that the inclusion of deliberate, reflective practice brings a greater actionable change, demonstrated through increased awareness and sit-

¹Robert Coles, "The Disparity between Intellect and Character," in *Ethics: The Essential Writings*, ed. Gordon Marino (New York: Random House, 2010), 350–55.

uational judgment applied to case studies. This chapter provides insight on the application of reflective practice to enhance learning.

A Leadership Course Taking Shape

The use of educational methods, by creating a dynamic learning environment, provides multiple opportunities for student success. While it is unrealistic to expect a perfect learning experience for every student, there are ways to structure a learning environment that is engaging and applicable for most students. By gauging the starting point of students, you can reach the greatest number through a creative learning environment. For example, at a more junior level of education, students may require more simplified instructions in smaller components during a longer period of time and with more interactive activities, whereas students at a more senior level of education are more likely able to understand requirements and will require less prodding to invest the effort to reach the outcomes. However, within the online learning environment, it is not just about the starting point of the student. In the traditional classroom setting, the instructor can easily guide students in the conversation or introduce new material as required versus online courses that require focused shaping and preparation of the material by ensuring proper placement and integration in the learning management system. In the online environment, preparation includes an understanding of the most basic learning concepts incorporated with multiple methods that creatively bring learning to life.

Whether teaching standard material across large cohorts of students or specializing in a specific subject for a small group, standards of learning follow the same general path: students are taught something, and then they apply what was learned. This is no difference in the online environment. However, online learning utilizes different methods to fit different outcomes, timelines, or populations, such as *replacing* face-to-face instruction or *enhancing* face-to-face instruction.² At the simplest level, courses

² Barbara Means et al., *Evaluation of Evidence-Based Practices in Online Learning: A Meta-Analysis and Review of Online Learning Studies* (Washington, DC: Office of Planning, Evaluation, and Policy Development, Department of Education, 2009), 3.

are constructed as stand-alone material completed through self-paced learning. In self-paced learning, the student has all the tools given up front, which are completed on their own without interaction. This works well for short training opportunities and skill development if the student can apply the material immediately through repetition, although there is the possibility for less retention.³ This method is not used for the online educational courses at the USAWC.

The AWC's Distance Education Program (DEP) uses a combination of methods to reach the students at varying levels based on the outcomes identified for each course. Rather than just identifying a meaningful outcome, the path to the outcome must be meaningful. The best way to create a useful learning environment for the online student is through multiple means of engagement, representation, and expression to “foster both attention and commitment,” “provide multiple pathways to build knowledge and comprehension,” and invoke “numerous methods to express their understanding.”⁴ In other words, the design of the course is well developed within the learning environment by making the best use of time, methods, and technology. Part of J. J. Walcutt and Sae Schatz's future learning ecosystem, the focus expands beyond opportunities for active engagement and checks-on-learning to verify attainment of the desired learning outcomes and ensure the educational structure supports learning opportunities.⁵

The DEP invests in a multimodal format with self-paced, individual application and live-virtual opportunities embedded in each course. A combination of self-paced learning with active learning occurs within asynchronous experiences, such as written forum posts shared with other students. This can be accomplished through discussion boards (forums), chat streams, or shared editing of written products. There are also

³ Peter C. Brown, Henry L. Roediger III, and Mark A. McDaniel, *Make It Stick: The Science of Successful Learning* (Cambridge, MA: Belknap Press, an imprint of Harvard University Press, 2014), 13.

⁴ Katie Novak and Catlin R. Tucker, *UDL and Blended Learning: Thriving in Flexible Learning Landscapes* (Oxford, UK: Impress Publishing, 2021), 35–39.

⁵ Brenda Bannan, Nada Dabbagh, and J. J. Walcutt, “Instructional Strategies for the Future,” in *Modernizing Learning: Building the Future Learning Ecosystem*, ed. J. J. Walcutt and Sae Schatz (Washington, DC: Government Publishing Office, 2019), 223–41.

synchronous requirements for students to complete, such as small group written products and exercises during the resident phases of the program. While students must progress through the program at the same pace and complete the same requirements, these varying degrees of synchronous and asynchronous opportunities keep the students engaged and actively learning the material for achievement of the program outcomes at a senior military education level. Students achieve through the attainment of passing grades for each assessment; however, the USAWC is actively pursuing ways to increase retention. One example is through reflection.

Stimulating Deeper Learning

Reflection in Learning

While the most basic learning concept forms the foundational structure of every course—learn something and then apply what was learned—reflection is one of the methods to support deeper learning for students. Reflection in learning is not a new concept. Many practitioners have espoused the importance of bringing judgment into learning, as early as Aristotle but also during the last century by John Dewey (1930s) and David A. Kolb (1970s) with experiential learning. According to Dewey, the first process in experiential learning is the experience of trial and error to arrive at learning, which works well when mastering a skill or requiring the same answer in every situation.⁶ Reflective activity is the second process of experiential learning, which provides deeper learning through problem-solving and the cognitive process of making connections between knowledge and experiences.⁷

Reflection, brought into the learning environment in a meaningful way, supports the connections for deeper learning. One key aspect of reflection is the deliberate process of thinking, often referred to as meta-

⁶ John Dewey, *Experience and Education*, Kappa Delta Pi Lecture Series (New York: Touchstone Book, an imprint of Simon & Schuster, 1938); and David A. Kolb, *Experiential Learning: Experience as the Source of Learning and Development* (Englewood Cliffs, NJ: Prentice-Hall, 1984).

⁷ David Boud, Rosemary Keogh, and David Walker, “What Is Reflection in Learning?” in *Reflection: Turning Experience into Learning*, ed. David Boud, Rosemary Keogh, and David Walker (New York: Routledge, 1985), 11–12.

cognition. Barry J. Zimmerman defines this as the “awareness of and knowledge about one’s own thinking.”⁸ Often, learning focuses on the content and maybe the application of the content. Expanding beyond this, reflection provides learners an opportunity to make meaningful connections allowing for “inquiry into one’s experience” by “converting experience(s) into meaningful learning.”⁹ Varying the learning activities initiates student knowledge transfer by making the connection between classroom material and applying that knowledge to other situations. This processing helps a student understand the breadth of the material beyond quick facts and helps process new information when faced with the experience in the future.¹⁰ Reflection, through the active process of thinking, leads to stronger learning when the knowledge is connected to experiences allowing for a rehearsal of how to change outcomes.¹¹ So, if reflection is a critical component to bridge learning to future actions, where has reflection been throughout education?

Reflection may be missing from most education since it is not as easy to assess. David Boud, Rosemary Keogh, and David Walker describe reflection in learning as “an active process of exploration and discovery which often leads to very unexpected outcomes.”¹² This is a different approach compared to the typical educational outcome. Instead of focusing on a specific outcome for an assessment, reflection allows for a more authentic learning environment supporting the processing of information just like a professional realistically makes decisions in their job.¹³ Students should be given time to understand and practice how the content applies to different situations. This occurs when learners purposefully reflect

⁸ Barry J. Zimmerman, “Becoming a Self-regulated Learner: An Overview,” *Theory into Practice* 41, no. 2 (Spring 2002): 65, https://doi.org/10.1207/s15430421tip4102_2.

⁹ Morris Fiddler and Catherine Marienau, “Developing Habits of Reflection for Meaningful Learning,” *New Directions for Adult and Continuing Education*, no. 118 (Summer 2008): 76, <https://doi.org/10.1002/ace.297>.

¹⁰ Flower Darby with James M. Lang, *Small Teaching Online: Applying Learning Science in Online Classes* (San Francisco, CA: Jossey-Bass, 2019), 181–82.

¹¹ Brown, Roediger, and McDaniel, *Make It Stick*, 27.

¹² Boud, Keogh, and Walker, “What Is Reflection in Learning?,” 7.

¹³ Joshua R. Eyler, *How Humans Learn: The Science and Stories Behind Effective College Teaching* (Morgantown: West Virginia University Press, 2018), 153.

outside of the classroom environment to arrive at a greater understanding of an experience.¹⁴

While not necessarily common across all educational institutions, reflection in learning can be found in professional military education. One example is the U.S. Coast Guard's Leadership Development Center, which has integrated daily reflective practice into the Midgrade Officer and Civilian Transition Course-1 to develop mental agility and intellectual curiosity.¹⁵ Similarly, the USAWC uses reflective practices within several courses. In particular, the USAWC Leadership Ethics Course provides students ample opportunity to make learning connections through reflective journaling.

Reflection in a USAWC Elective Course

As adult learners with an average of 20 years of professional service, students attending the USAWC already have multiple experiences to draw from. The online ethics course builds on their service by guiding the learning to new information and motivating the students to make connections to their experiences. The course provides the same basic concept for each lesson: learn about a theory or concept; understand how the material applies to a specific situation (case study); use critical thinking to craft a written response; and receive a grade—ideally passing the requirement—while moving on to the next lesson. Within this construct there are assumptions. First, it is assumed that students take ownership in their learning. As in constructivism, individuals will construct their own knowledge in addition to what is provided.¹⁶ This begins when, as an elective course, the student made a choice to learn more on the subject. In this course, construction began with identifying outcomes and assessments

¹⁴ Olenda E. Johnson, "Creating the Space to Think: The What, Why, and How of Deliberate Reflection for Effective Leadership," *Journal of Character and Leadership Development* 7, no. 1 (Winter 2020): 20–32.

¹⁵ Tom Miller, Jonathan Tice, and Tommy Brabson, "Metacognitive Reflection: The Framework for Facilitating Reflective Practice during the Coast Guard Midgrade Officer and Civilian Transition Course," *Journal of Military Learning* (October 2021): 62.

¹⁶ Robert D. Tennyson and Ashley Volk, "Learning Theories and Educational Paradigms," in *International Encyclopedia of the Social and Behavioral Sciences*, 2d ed., ed. James D. Wright (Amsterdam: Elsevier, 2015), 699–711, <https://doi.org/10.1016/B978-0-08-097086-8.92036-1>.

based on the content that guides the student to the mandatory learning objectives. As an elective, students that are interested in expanding the basic learning are provided focus questions to assist with constructing their own learning experience, which is done through the reflective process. The choice then falls on the student for ownership in their learning by deciding whether to complete the minimum requirements or to expand into optional reflective opportunities.

If the student takes ownership in their learning, the second assumption expects the student to understand how the information applies to other situations, beyond the initial case study and requirement. In this assumption, students take it on themselves to ponder how the information could have helped in a past situation they experienced, or how it could help in a current or future situation. Because of uncertainty and risk in the military, this assumption lends itself to the primary reason for offering the course, which supports leaders learning how to effectively adapt and use judgment in a constantly changing environment.¹⁷ The course provides a realistic, though safe, opportunity to wrestle with ethical challenges through the given case studies. Students then must understand how their learning transfers to other situations.

Adding to the second assumption, the third assumption expects that the students will retain the information for future use. Just as they may think about the information beyond the given case study, students must make the leap to apply what they learn when faced with challenges as a leader. Students reflecting on potential follow-on application are “crystallizing and reinforcing previous learning, developing concepts and generalizations for future use.”¹⁸ Similar to constructing their outcomes, students will experience more retention when they make the connections through reflection.

Reflective Journaling

The literature on reflective practice and these three assumptions led to

¹⁷ Jan Fook and Fiona Gardner, *Practising Critical Reflection: A Resource Handbook* (Berkshire, UK: Open University Press, an imprint of McGraw-Hill, 2007), 10.

¹⁸ Boud, Keogh, and Walker, “What Is Reflection in Learning?” 13.

the addition of a reflective journal requirement in the course as the catalyst for stimulating deeper learning. But why journaling? If reflection is purposeful, deliberate thinking, why must it be written? The Center for Creative Leadership supports the reflective practice of journaling in leadership programs because it allows for personal integration “relevant to their own situations” and improves the retention for “transfer back to their home environment” once the formal learning program ends.¹⁹ Just like online learning should continuously draw the student back to the outcomes through an integration of concepts, journaling provides an opportunity for the student to revisit the reflections to make connections between education and experience.

Privacy is another key aspect of written reflective journals as compared to other learning requirements. Within the online environment, a common assessment requires students to share written work through asynchronous forums, which are meant to mimic in-class verbal discussion. Knowing that other students are reading your work, the typical student will shy away from writing anything that would bring negative critique. Conversely, journaling in an online course provides a private outlet for written thought that is shared only between the student and instructor. Learning management systems with a journal feature allow for privacy in this way; emailing the journal entry from student to professor is a secondary option to ensure privacy. Providing this additional component beyond the initial assessment encourages students to share their feelings in a nonthreatening way, while also providing a check on learning.

A tip given to the students is to not complete the journal in one sitting. Students are encouraged to copy the questions into a word document and complete an initial response to the reflection prompts. They are then encouraged to set it aside and return the next day or a few days later to read their answers and reflect on any additional thoughts that arise. While most students prefer to complete the assignment and turn it in for grading immediately, a few students will use this method. Indications that this has happened appear in statements such as “when I first answered this

¹⁹ Talula Cartwright, *Developing Your Intuition: A Guide to Reflective Practice* (Greensboro, NC: Center for Creative Leadership, 2004), 27.

I . . .” or “the next day it occurred to me . . .”²⁰ Those are also the students that tend to write longer journal entries and appear to put more thought into the overall requirement, which demonstrates metacognition.

Testing the Options for Reflection

In 2022, the first iteration of the USAWC Leadership Ethics Course was offered to DEP students. Being strictly online, the elective course contained weekly lessons (Monday through Sunday) with each lesson requiring asynchronous written posts in the discussion forum. Although the student could conduct reflection during the weekly discussion thread, there was a tendency to focus only on an initial application of the material to the case study, which is how they will pass the class. Reflection was added into the first iteration of the course through journals as optional weekly opportunities by providing a series of focus questions based on the main topics and learning material. This approach garnered modest support, with one-third of students experimenting with the journal at some point during the course, and approximately 20 percent of the students completing every weekly journal.

Student feedback indicated that having the option to journal was a refreshing opportunity. As one student stated, “I was amazed that the questionnaire and journal we completed during the first week popped into my head while going through the reading this week. I really see that I have a greater understanding of how viewpoints differ when it comes to making decisions.”²¹ Another student commented that it was a good option to try but preferred to use the reflective questions to focus study preparation rather than taking the time/energy to write. Depending on the course length and amount/timing of the work required, optional journaling may be the best choice with the understanding that it is likely to receive a lesser focus from students. Conversely, mandatory journal entries are ideal to support the use of reflection in learning.

There are several ways to incorporate mandatory journaling. The second and third iterations required journals at the beginning and end of the

²⁰ Student journal entries, 2022–23.

²¹ Student journal entry, 11 November 2022.

course, with optional reflective questions identified during the central weeks of the asynchronous discussions. Aside from the mandatory journal entries, this method received a slightly higher level of interest as the previous course with purely optional journaling. During the second iteration, approximately one-half of students experimented and approximately 30 percent maintained weekly journaling. The third iteration of the course, which also included mandatory journals at the beginning and end, maintained 36 percent of students completing all the optional journaling.

Based on the usage, the optional journaling did not meet the desired outcome for reflection in learning, but it did provide insight for extending the use of journaling:

1. Students were given the opportunity to receive extra credit if they completed optional journals, which provided an incentive for them to explore the opportunity.
2. When considering the assessment for the mandatory journals, grading can be simplified as pass-fail; as long as students answer the question(s), they receive full credit.
3. Journals are meant for reflection and personal thought, so students are encouraged not to reference any material.
4. Students are told that the typical written requirements, such as correct grammar or word count, are not graded in the journal.

This openness in the grading and structure provides more freedom for the students to not feel pressure to perform in the journals as they do in other writing requirements. Even though most students did not complete journals throughout, some students clearly benefited with the added optional journaling. As one student identified:

I didn't think much about the journal entry we did during lesson 4, other than to answer the question you asked. But when I was writing my post the following week I thought "That's what it was about." I now see how I can apply what I learned, even though I didn't see it at the time.²²

²² Student feedback to author, 8 December 2022.

In each iteration of the course, either purely optional or limited mandatory requirements, the use of reflective journaling demonstrated that students made deeper connections with the material.

The Challenge

In each iteration of the online elective, students consistently demonstrated the ability to apply the concepts to case studies. Based on the interest of students through a high selection rate of the elective course, a short seminar option was provided for students during their resident phase. Similar to the online course, students must elect the short course, so ideally every student has interest in the topic and wants to put forth the effort to improve their ethical decision making. In this version of the course, the same primary topics are covered, but they occur during two days through in-class instruction, case study review in small groups, and problem-solving discussions.

In the first seminar version, the students completed the requirements as expected—they made the linkages between the material and the case studies—with one exception. As a natural phenomenon with classroom discussion, the conversations consistently veered into current professional dilemmas and leadership difficulties. By questioning the underlying issues, it became apparent that the students were able to identify the correct actions in complex situations, like the written communication in the asynchronous forums; however, the students stated that they would not necessarily take those actions if faced with the situation.

After further review of the first iteration online, student journals also identified a disparity between identifying the correct solutions and applying those actions if confronted with a similar situation. This incongruency led to the introduction of the two mandatory journals in the second iteration of the online course. And just like the previous year, the disparity was once again noted during both the online course and the seminar resident course. In other words, the critical thinking demonstration was on target with the expected outcome in the written assignment—each person took the appropriate steps in solving or actioning the ethical problem—but they admitted that their actions would be different if faced with the situation. Further investigation identified that the students were wrestling

with the case study based on the information given. Students admitted that they understood what was required in the written forum, but that they would take more time and gather more information before acting. It was now clear that there must be a greater emphasis on reflective journaling to create an increase in the learning, both through understanding the disparity and for further leader development.

Changing the Perspective on Student Learning

The online course already included asynchronous application, some reflective journaling, and optional live-virtual discussions. It was evident that adding a new perspective through reflective journaling was the best method to create the learning connections for an online course. This includes both increasing the required number of mandatory journals as well as rethinking the structure of the journals.

To make the journals more effective in student learning, and begin closing the gap between intellect and character, the journals are purposively added to each lesson with three levels of reflection expanding on the lesson topic. Purposively adding journals means that the journal entry is not freeform, or that the student chooses a question out of a series of options. Journals purposively added to each lesson expound on a key element, typically with an activity centered around the lesson's theme. In this course, the activity focuses on an exercise, short questionnaire, or article. Then three questions or prompts, based on the activity, guide the reflection for the journal entry.

Why three questions for the journal? Three prompts move the student through increasing cognitive levels: understand, analyze, and create.

Understand

The first question focuses on individual awareness and thinking about the material from a personal perspective. Students are encouraged to interpret what they learned and explain ideas in their own words.²³ Typical questions are: What were your first thoughts when you saw the results

²³ Patricia Armstrong, *Bloom's Taxonomy* (Nashville, TN: Vanderbilt University Center for Teaching, 2010).

[of the questionnaire you completed]? Were you surprised by your results; explain why or why not?

Analyze

The second question asks the student to think more deeply about what they learned through a level of analysis that “ask[s] the learner to organize elements within a structure, distinguish relevant from irrelevant information, or deconstruct underlying values and biases.”²⁴ This prompt focuses on the connection with other situations, like examining the information through a different lens or comparing/contrasting the concept. A typical prompt is: think of a time when you have seen this occur in a previous situation; describe what happened.

Create

The third question requires forethought in how it could apply in the future. Students are prompted to construct a new framework, formulate a way to incorporate their knowledge when developing others, or conceptualize a different end state to a problem.²⁵ Typical questions are: Now that you know your result, how will you use this as a leader in the future? How can you use this information to develop those in your unit?

By guiding the reflection in three separate prompts, students can think about the information in a way that they might not naturally consider when just thinking about the material. This purposeful injection of mandatory journals, based on the three cognitive levels of reflection, provides an opportunity for closing the gap between thought and action. Initial feedback from students indicates that reflection creates an actionable change in the disparity between intellect and character. For example:

A case study given to students identifies the potential misconduct of a mentor and superior officer. Students are asked what they would do based on the informa-

²⁴ Toyin S. Tofade, Jamie Elsner, and Stuart T. Haines, “Best Practice Strategies for Effective Use of Questions as a Teaching Tool,” *American Journal of Pharmaceutical Education* 77, no. 7 (2015): 2, <https://doi.org/10.5688/ajpe777155>.

²⁵ Tofade, Elsner, and Haines, “Best Practice Strategies for Effective Use of Questions as a Teaching Tool,” 3.

tion given. The written discussion forum elicited the same reaction to the situation—confront the individual—from nearly every student. Typically, at least one student would identify that confrontation would not be required but acknowledged that they may confront the person after finding out more information. Conversely, when students were given the opportunity to reflect on their actions in a private journal with the three increasing cognitive prompts, every student identified a deliberate process of finding out more information, eliciting options from others, having another person present when confronting the individual, and only then would the students confront the individual. This reflective journaling allows for a safe way to practice deliberate decision making and build the knowledge for future use.

Effective Use of Reflection

A fourth iteration of the course concluded with mandatory reflective journaling throughout vice the optional journaling. The addition of the mandatory journals strives for metacognition. The use of reflective journaling supports character development for leaders through a deeper level of learning and increased integration of thought and action. Student feedback indicates an appreciation of the extra opportunity to work through the information in a private journal, with 100 percent of students indicating the journals were an effective tool in support of their learning.²⁶ One student commented that the journals gave the opportunity for “thinking about the material and scenarios with no format, a free flow of thought, but giving time to think about it and reflect back on it,” which was appreciated to expand on the asynchronous forum posts that everyone sees.²⁷ Another student also liked the privacy, stating “I’m really glad that other people couldn’t see what I wrote in the journals because it was more for me than it was for anybody else. They were really invaluable because it

²⁶ Student and survey feedback, May 2024.

²⁷ Student A feedback to author, 7 May 2024.

kind of tied our own thoughts back to the lesson, and I saved them so I can go back to them later.”²⁸ Students also saw the usefulness of reflection skills for after their educational program. Feedback from one student identified that they would continue professional development after the learning experience: “As I conclude this course, I plan to continue integrating some of the authors and readings into my leadership tool kit. I’ll also continue using reflective journaling as a way to help me stay on track and monitor my progress.”²⁹ Another student identified the usefulness for their professional and personal development:

Something I didn’t think I’d like is the journaling. But then reading the material and actually thinking about it and in a different way really forced me a little bit out of my comfort zone. I have never been a journaling person, but after the first lesson, when you introduced journaling to us, I bought myself a journal. It’s just a sentence a day, but it’s really changed me. It’s a nice way to wake up and start the day.³⁰

Future Directions

The inclusion of reflection, typically through a written method such as journaling, can be applied to a variety of courses. The use of journals in online courses provides the best opportunity for students to conduct an additional assessment of the material, while also allowing for introspection beyond the written online forum. For courses employing live-virtual or in-person discussions, the reflective journal provides additional time for analyzing the learning material versus on-the-spot recollection. During course creation, instructors must identify whether mandatory or optional journaling would elicit the desired outcomes, especially if the focus is self-development versus an additional assessment of learning.

While classroom discussions allow for shaping any disparity between intellect and character, the online environment must be carefully created

²⁸ Student B feedback to author, 7 May 2024.

²⁹ Student C feedback to author, 7 May 2024.

³⁰ Student D feedback to author, 7 May 2024.

to cultivate similar shaping. Mandatory written journals support the basis of this work—students who reflectively digest the material through journaling reach a deeper understanding and complete the three assumptions: take ownership, understand other applications, and retain. While journaling provides the baseline for student achievement beyond the minimum work required, a fourth assumption was added to further solidify learning. The final assumption is that by adding a mandatory reflective journal for every lesson, the students take their learning to the next level through enhanced understanding, analysis, and creation, which are the three aspects to the required journals.

Verbal feedback from students and surveys at the end of each course supports the inferences made in this chapter. For example:

Even though we haven't had a chance to meet in person, I wanted to take a moment to say thank you for teaching the Leadership Ethics course last year. Since taking your class, I've been journaling every day for almost a year now (who knew I had that kind of commitment in me?) and it's led me down a path of positive habits. Your course really inspired me to broaden my reading list, too. I've been devouring books with similar concepts ever since. I have to tell you that your course had a major impact on my personal journey over the past year. It's helped me be more empathetic, aware, and mindful. I've been more focused on personal improvement, introspection, and reflection—and all of this has led to an overall more positive life perspective. So, I just wanted to say thank you for giving me that starting point.³¹

Learning about the content may achieve the expected outcome of a course; however, the student receives a much larger reward when their educational experience continues beyond the classroom. Using reflective practice is just one method to support classroom achievement, with the belief that students will extend the practice to lifelong learning.

³¹ Email message from student to author, 20 March 2023.



CHAPTER 6 ●

Localized Professional Development

Creating Base-Wide Solutions to a Military-Wide Problem

Layton Graves, EdD

Introduction

A quantitative assessment utilizing Roger Boshier's Education Participation Scale was conducted at Freedom Air Force Base (FAFB, a pseudonym to adhere with U.S. Department of the Air Force research protocols).¹ The findings indicated multiple motivations for airmen to pursue voluntary education and suggest that airmen would attend professional development seminars targeted at achieving objectives beyond professional development.² Initiatives at the base level are best poised to align programs to local needs/objectives. This study was conducted on a U.S. Air Force (USAF) installation; however, the findings may have broader applicability to the Services as a whole. Additional research is recommended to explore this concept in greater detail.

Professional development (PD) is widely recognized as beneficial to

¹ Roger Boshier, "Psychometric Properties of the Alternative Form of the Education Participation Scale," *Adult Education Quarterly* 41, no. 3 (1991): 150, <https://doi.org/10.1177/0001848191041003002>.

² Within the Services, airmen is considered a gender-neutral term for Air Force service-members.

both employees and organizations.³ The motivations for an organization to implement PD and for an individual to pursue PD are important to understand when developing optimal programs.⁴ Additionally, perceptions of PD programs can drastically influence program utilization and the application of course material.⁵ The military presents a unique context for PD with motivations and barriers that are distinct to the lifestyle of servicemembers and their families.⁶ Specifically, within the USAF, multiple challenges for PD programs are visible. Potential barriers include personnel working at more than 99 geographically separated bases, a population who generally distrusts institutional sponsored remote learning, and shifts in organizational priorities from a leadership team that changes every four years.⁷ The extent to which institutional practices affect PD utilization in society is well researched; however, very few studies exist on how the unique context of the USAF affects PD program optimization

³ Audra Classen, Eva Horn, and Susan Palmer, "Needs of Military Families: Family and Educator Perspective," *Journal of Early Intervention* 41, no. 3 (2019): 246–47, <https://doi.org/10.1177/1053815119847235>; Insu Kim and Jo Won-kwang, "The Effects of Joint Education Programs on Cohesion among South Korean Military Academy Cadets," *Korean Journal of Defense Analysis* 26, no. 4 (2014): 517–18; and Jeremy Jordan and Christopher Curtis, "Evaluating the Impacts of Technology Education on Military Maintenance Students," *Journal of Stem Teacher Education* 47, no. 2 (2010): 41–42.

⁴ Boshier, "Psychometric Properties of the Alternative Form of the Education Participation Scale," 150–51.

⁵ Ted Futris et al., "The Impact of Organizational Support on the Transfer of Learning to Practice," *Children and Youth Services Review* 51 (2015): 41–42, <https://doi.org/10.1016/j.childyouth.2015.01.019>.

⁶ Classen, Horn, and Palmer, "Needs of Military Families," 233; Jennifer M. D. Kremkow and Erinn H. Finke, "The Effectiveness of a Distance Peer Mentor Training Program with Military Spouses with Children with Autism," *Journal of Autism and Developmental Disorders* 50, no. 4 (2020): 1097, <https://doi.org/10.1007/s10803-019-04334-0>; and Alexa Smith-Osborne, "An Intensive Continuing Education Initiative to Train Social Workers for Military Social Work Practice," *Journal of Social Work Education* 51, no. suppl (2015): 89–90, <https://doi.org/10.1080/10437797.2015.1001290>.

⁷ "Locations," AirForce.com, accessed 18 February 2024; Jason Newcomer et al., "Just Checking the Box: Do Our Airmen Value Their CCAF Degree," *Air & Space Power Journal* 30, no. 1 (2016): 39–42; and "Military Service Chiefs," Joint Chiefs of Staff, accessed 18 February 2024.

or the extent to which the motivations and barriers for PD utilization for airmen mirror overarching trends.⁸

Conceptual Framework

The Education Participation Scale (EPS) has an extensive history as an effective tool for measuring and understanding the motivations of adult learners.⁹ The EPS identifies seven distinct motivations for adult learners to pursue voluntary education: (a) communication improvement, (b) social contact, (c) education preparation, (d) professional advancement, (e) family together, (f) social stimulation, and (g) cognitive interest. It was originally developed by Roger Boshier through a study of 845 adult learners and is a reliable measure validated through a large base of empirical research.¹⁰ The U.S. military has a demonstrated history of using PD programs that align with the seven motivations for voluntary education expressed in the EPS framework.

Literature Synthesis

Cognitive Interest

PD programs can target adult learners who are motivated to pursue education due to cognitive interest—seeking specific knowledge or learning for the sake of learning itself.¹¹ PD as a means of covering a specific topic of interest is often used within the military, especially within the medical corps. Military populations are known to have unique medical needs due

⁸ Lisa A. Boyce, Stephen J. Zaccaro, and Michelle Zazanis Wisecarver, “Propensity for Self-Development of Leadership Attributes: Understanding, Predicting, and Supporting Performance of Leader Self-Development,” *Leadership Quarterly* 21, no. 1 (2010): 160, <https://doi.org/10.1016/j.leaqua.2009.10.012>; and Kathryn Hayes et al., “Beyond Funding: How Organizational Resources Support Science Professional Learning,” *American Journal of Education* 126, no. 3 (2020): 389–422, <https://doi.org/10.1086/708266>; Emily Grace Keenlyside, “Critically Engaging Volunteer Guides: A Study of Adult Learning in an Art Museum,” *Canadian Journal for the Study of Adult Education* 31, no. 2 (2019): 8–11.

⁹ Boshier, “Psychometric Properties of the Alternative Form of the Education Participation Scale,” 150–67.

¹⁰ Boshier, “Psychometric Properties of the Alternative Form of the Education Participation Scale,” 151–53.

¹¹ Boshier, “Psychometric Properties of the Alternative Form of the Education Participation Scale,” 156.

to military-specific occupational hazards and frequent changes in medical providers as a result of assignment changes.¹² A study of 268 USAF and Department of Veterans Affairs (VA) social workers who attended 10 PD workshops addressing risk, resilience, coping strategies, and social support for military populations showed an increase in self-efficacy for 90 percent of the participants and an increase in knowledge of how to serve military populations for 96 percent of the participants.¹³ Additionally, a similar study was conducted on military endocrinologists who attended a three-day PD workshop on diabetes care for military populations.¹⁴ Attending the PD session resulted in an increase in short-term knowledge retention on military diabetes care for 90 percent of the participants.¹⁵ In both cases, PD was effectively used to hone skillsets valuable to the military through courses that incorporated curriculum designed to address a topic of interest.

Professional Advancement

PD programs can target adult learners who are motivated to pursue education out of a desire for professional advancement—improving status in an existing job or preparing for a new job.¹⁶ One of the largest, and arguably most successful, professional advancement and career assistance programs in the world exists within the U.S. military.¹⁷ The VA Transition Assistance Program (TAP) is a congressionally mandated series of PD seminars and career coaching that serves more than 200,000 personnel

¹² Classen, Horn, and Palmer, “Needs of Military Families,” 233.

¹³ Smith-Osborne, “An Intensive Continuing Education Initiative to Train Social Workers for Military Social Work Practice,” 95–100.

¹⁴ Darrick Beckman et al., “Evaluation of an Interprofessional Continuing Professional Development Course on Comprehensive Diabetes Care: A Mixed-Methods Approach,” *Journal of Evaluation in Clinical Practice* 25, no. 9 (October 2018): 149–50, <https://doi.org/10.1111/jep.13033>.

¹⁵ Beckman et al., “Evaluation of an Interprofessional Continuing Professional Development Course on Comprehensive Diabetes Care,” 151.

¹⁶ Boshier, “Psychometric Properties of the Alternative Form of the Education Participation Scale,” 156.

¹⁷ Xiaoxue Li, “Improving the Labor Market Outcomes of US Veterans: The Long-Run Effect of the Transition Assistance Program,” *Defence and Peace Economics* 31, no. 1 (2020): 48, <https://doi.org/10.1080/10242694.2018.1532229>.

separating from the military each year. TAP was implemented in 1991 and has garnered praise for fostering a 10-percent increase in full-time worker status and a 17-percent increase in the use of government-sponsored benefit programs for recently separated veterans.¹⁸

Furthermore, PD programs outside of TAP are demonstrated to help military personnel professionally advance. An experimental career transition program was tested within the U.S. Navy in 2014 with a high rate of success.¹⁹ The study involved 41 sailors stationed in Naples, Italy, who indicated an intent to separate from the military. All participants were provided with group training and individualized career coaching sessions. Three months after the intervention, 100 percent the participants were employed full time—a success compared to the average employment rate of 91 percent for personnel who separate from the Navy.²⁰ PD is often used to aid military personnel in career transitions and advancements with a significant level of success.

Family Together

PD programs can target adult learners who are motivated to pursue education as a means of bringing their family together—bridging generational gaps and improving family relations.²¹ The U.S. military recognized the benefit PD can have on improving family relations and created a program designed to help military families who are affected by autism. The Military Spouse Online Autism Relocation Readiness (MilSOARR) program is a virtual community with an emphasis on mutual support and educational resources.²² MilSOARR is targeted at decreasing stress and anxiety during military moves for families with a child on the autism spectrum

¹⁸ Li, “Improving the Labor Market Outcomes of US Veterans,” 48–52.

¹⁹ Kathryn Troutman and John Gagnon, “The Role of Professional Coaching and Resume Writing in Successful Veteran Transitions,” *Career Planning & Adult Development Journal* 30, no. 3 (2014): 210.

²⁰ Troutman and Gagnon, “The Role of Professional Coaching and Resume Writing in Successful Veteran Transitions,” 212–13.

²¹ Boshier, “Psychometric Properties of the Alternative Form of the Education Participation Scale,” 156.

²² Kremkow and Finke, “The Effectiveness of a Distance Peer Mentor Training Program with Military Spouses with Children with Autism,” 1099.

and connecting families with mentors who have previously overcome similar challenges. The virtual training and connection platform is still being analyzed for program outcome trends; however, many military spouses within the program report high levels of satisfaction and subsequently volunteer to become mentors for new program participants.²³

Social Contact

PD programs can target adult learners who are motivated to pursue education as a means of social contact—meeting new people, networking, and making friends.²⁴ An increase in the frequency of military networking can have significant benefits on unit cohesion and morale.²⁵

The South Korean military, which has an organizational structure modeled on the U.S. military, conducted a mixed-methods study on military cadets who attended a training course with cadets of different branches.²⁶ Through observations and self-assessment questionnaires, it was determined that cadets who attended joint service training produced a more collaborative atmosphere that extended outside of the classroom, shared a greater level of course related information with their peers, and obtained a notably higher level of camaraderie compared to cadets who did not attend the joint service training.²⁷ Educational courses can foster an environment that leads to an increased level of cross-talk and social contact between military personnel.

Social Stimulation

PD programs can target adult learners who are motivated to pursue education as a means of social stimulation—providing a break from the day-

²³ Kremkow and Finke, “The Effectiveness of a Distance Peer Mentor Training Program with Military Spouses with Children with Autism,” 1106–7.

²⁴ Boshier, “Psychometric Properties of the Alternative Form of the Education Participation Scale,” 156.

²⁵ Kim and Won-kwang, “The Effects of Joint Education Programs on Cohesion among South Korean Military Academy Cadets,” 509–11.

²⁶ Kim and Won-kwang, “The Effects of Joint Education Programs on Cohesion among South Korean Military Academy Cadets,” 505–8.

²⁷ Kim and Won-kwang, “The Effects of Joint Education Programs on Cohesion among South Korean Military Academy Cadets,” 514–18.

to-day work routine.²⁸ Social stimulation is noted as one of the primary needs of airmen following a traumatic incident, such as an occurrence of death by suicide within their unit.²⁹ The USAF created a PD program targeted at increasing social stimulation—the Wingman-Connect Program.³⁰ This program provides airmen an opportunity to take a break from their normal duties and connect with others during their standard work hours. A mixed-methods study of 1,485 airmen across 215 classrooms showed a correlation between participation in the program and a decrease in self-reported feelings of isolation.³¹ Education as a form of social stimulation can be used to connect airmen and reduce feelings of isolation.

Education Preparation

PD programs can target adult learners who are motivated to pursue education as a means of education preparation—remediating educational deficiencies or preparing for a higher level or more specialized education.³² The U.S. military uses Basic Military Training (BMT) as a form of education preparation. During BMT, recruits from a wide variety of backgrounds and experiences receive foundational knowledge that is built on in technical school courses and subsequent professional military education (PME). Additionally, BMT is demonstrated to be an ideal time to remediate educational deficiencies on a wide variety of topics, to include the repercussions of tobacco use.³³ In 2020, more than 2,000 airmen

²⁸ Boshier, “Psychometric Properties of the Alternative Form of the Education Participation Scale,” 156.

²⁹ Peter A. Wyman et al., “Wingman-Connect Program Increases Social Integration for Air Force Personnel at Elevated Suicide Risk: Social Network Analysis of a Cluster RCT,” *Social Science & Medicine* 296 (2022): 1–2, <https://doi.org/10.1016/j.socscimed.2022.114737>.

³⁰ Wyman et al., “Wingman-Connect Program Increases Social Integration for Air Force Personnel at Elevated Suicide Risk,” 2.

³¹ Wyman et al., “Wingman-Connect Program Increases Social Integration for Air Force Personnel at Elevated Suicide Risk,” 2–5.

³² Boshier, “Psychometric Properties of the Alternative Form of the Education Participation Scale,” 156.

³³ Melissa Little et al., “Factors Associated with Cigarette Use during Airmen’s First Year of Service in the United States Air Force,” *Military Medicine* 185, nos. 1–2 (January–February 2020): 213, <https://doi.org/10.1093/milmed/usz155>.

voluntarily participated in a tobacco-free PD program during BMT and demonstrated lower reported rates of tobacco usage after leaving BMT and throughout technical training programs compared to peers who did not participate in the PD program.³⁴ PD is an effective tool for rectifying previous educational deficiencies and preparing personnel for future educational opportunities.

Communication Improvement

PD programs can target adult learners who are motivated to pursue education as a means of communication improvement—enhancing skills outside of the stated course material, typically verbal and written skills.³⁵ The USAF successfully uses education for communication improvement during Squadron Officer School (SOS) at Air University (AU), a mandatory PME course for officers with the rank/grade of captain/O-3.³⁶ During SOS, officers are instructed on leadership principles and effective stress coping strategies through coursework, case studies, and role playing. Officers scored higher on leadership assessments at the conclusion of SOS compared to leadership assessments prior to SOS.³⁷

Additionally, communication improvement is a frequent learning objective of PME. Both guided reflection and group reflection are demonstrated to be present within PME and are proven to lead to a diversity of ideas among participants.³⁸ A study was conducted on 498 personnel with the rank/grade of captain/O-3, where an intervention group received a 30-minute seminar on leadership and a 15-minute seminar on how to navigate leadership resources available to officers.³⁹ The control group

³⁴ Little et al., “Factors Associated with Cigarette Use during Airmen’s First Year of Service in the United States Air Force,” 215–19.

³⁵ Boshier, “Psychometric Properties of the Alternative Form of the Education Participation Scale,” 156.

³⁶ Gene M. Alarcon et al., “Leadership and Coping among Air Force Officers,” *Military Psychology* 24, no. 1 (2012): 36.

³⁷ Alarcon et al., “Leadership and Coping among Air Force Officers,” 38–43.

³⁸ Tim Sparks, “What Do Creative Problem Solvers Need,” *Military Scholarship on Teaching and Learning Forum*, 2023, video, 58:50.

³⁹ Boyce, Zaccaro, and Wisecarver, “Propensity for Self-Development of Leadership Attributes,” 166.

watched 45-minutes of cartoons. All participants were asked to complete surveys designed to measure motivation and propensity to leadership self-development. The study uncovered that the leadership PD had a notable impact on those in attendance and measurably increased their propensity for future engagement in self-directed learning.⁴⁰ In each study, military PD was used as an effective tool for increasing communication improvement.

Implementation Barriers

Despite the demonstrated success of PD within the military, there are multiple barriers hindering the effective utilization of PD within the USAF. There is ambiguity as to the importance of pursuing PD by formal requirements and unwritten rules surrounding USAF education.⁴¹ Self-improvement is a factor considered on annual performance reports, leading some supervisors to encourage their airmen to attend PD out of a desire for career advancement. An inconsistent standard to pursue PD can be present depending on how an individual supervisor perceives the importance of having education listed on a performance report. A research study on airmen who recently completed the requirements for an associate degree through the Community College of the Air Force (CCAF) in Birmingham, Alabama, uncovered a perception among some graduates that the degree did not hold value outside of the USAF and was merely a box that needed to be checked to be competitive for a promotion.⁴² Although it is unclear whether the perception of external value for a CCAF degree matches the reality, a negative perception alone can lead to low enrollment and suboptimal learning outcomes.⁴³

A perception of low organizational support or of inferior quality can impact both enrollment and knowledge retention regardless of wheth-

⁴⁰ Boyce, Zaccaro, and Wisecarver, "Propensity for Self-Development of Leadership Attributes," 168–74.

⁴¹ Newcomer et al., "Just Checking the Box," 32.

⁴² Newcomer et al., "Just Checking the Box," 30–45.

⁴³ Newcomer et al., "Just Checking the Box," 41–42.

er that perception is justified.⁴⁴ A 2019 study of adult learners concluded that individuals who had a negative perception of the PD seminar prior to entering the course did not experience the long-term benefits in knowledge retention or self-efficacy noted in those who had a positive perception of the PD seminar prior to entering.⁴⁵ It is imperative that airmen perceive a PD program as beneficial to achieve the desired outcomes of the program.⁴⁶

Empirical Study

The USAF needs to examine and understand the motivations of airmen to pursue PD and their perception of USAF PD as a central part of the effort to increase program effectiveness. The literature suggests that the motivations of adult learners within a voluntary educational program will include both motivations that align with stated course objectives and motivations that align with personal interests.⁴⁷ It is possible that the motivations of airmen to pursue PD at FAFB could differ from the hypothesized results. Therefore, all seven of Boshier's motivations to pursue education will be examined.

Context

This study takes place at FAFB in North America—home to around 10,000 military personnel, contractors, family members, and retirees. Approximately 3,000 personnel at FAFB are active-duty military, with an increase in the number of assigned personnel notably growing each year. At any given time, an undisclosed portion of the active-duty population

⁴⁴ Kyriaki Makopoulou et al., “An Investigation into the Effects of Short-Course Professional Development on Teachers’ and Teaching Assistants’ Self-Efficacy,” *Professional Development in Education* 47, no. 5 (2021): 790–91, <https://doi.org/10.1080/19415257.2019.1665572>; and Newcomer et al., “Just Checking the Box,” 41–42.

⁴⁵ Makopoulou et al., “An Investigation into the Effects of Short-Course Professional Development on Teachers’ and Teaching Assistants’ Self-Efficacy,” 790–91.

⁴⁶ Makopoulou et al., “An Investigation into the Effects of Short-Course Professional Development on Teachers’ and Teaching Assistants’ Self-Efficacy,” 790–91; and Newcomer et al., “Just Checking the Box,” 41–42.

⁴⁷ Adrian Blunt and Baiyin Yang, “Factor Structure of the Adult Attitudes Toward Adult and Continuing Education Scale and Its Capacity to Predict Participation Behavior: Evidence for Adoption of a Revised Scale,” *Adult Education Quarterly* 52, no. 4 (2002): 299–301.

is deployed or temporarily stationed at another geographic location resulting in a rough estimate of about 2,000 active-duty airmen physically present at FAFB.

The education center at FAFB holds a weekly briefing on the Air Force Tuition Assistance (AF-TA) program. Each seminar is scheduled for one hour and is a prerequisite to participating in the AF-TA program. AF-TA is a Service-sponsored program where eligible airmen can receive up to 100 percent of the cost of tuition (up to \$250 per semester hour) for the pursuit of an accredited college degree without consuming any post-9/11 GI Bill entitlements.⁴⁸ This program is only available to active-duty airmen or airmen employed in a full-time status within the Air National Guard or Air Force Reserves. The AF-TA entitlement is provided by the USAF directly to the higher education institution and is contingent on: (a) initial approval from the airman's direct supervisor, (b) an approved degree plan from the education center, (c) satisfactory performance reports and fitness reports, and (d) successful course completion with at least a C grade for undergraduate courses or a B for graduate courses.⁴⁹ The AF-TA briefing at FAFB covers all criteria for obtaining degree plan approval and AF-TA funding for both enlisted personnel and commissioned officers. The AF-TA briefing at FAFB has an average attendance of 15 airmen each week. This PD course can be viewed as targeting the education motivation of education preparation due to its stated focus of preparing airmen to use a service designed to fund additional educational pursuits.

Method

This quantitative research project involved the voluntary collection of data from airmen at FAFB regarding their motivations to pursue voluntary PD. There is currently no existing data on the motivations of airmen to pursue USAF sponsored PD at FAFB or any specific USAF installation. The research design, variables, and data collection methods are based on

⁴⁸ "Military Tuition Assistance Program," Air Force Personnel Center, accessed 18 February 2024.

⁴⁹ "Military Tuition Assistance Program."

previous studies that effectively used the EPS to measure the motivations of adult learners.⁵⁰ Additionally, this research project is constrained by USAF regulations limiting data that can be collected by students for the partial fulfillment of an academic degree.⁵¹ Future studies with a broader scope should be conducted to validate the results and to identify potential correlations across multiple USAF installations.

Research Questions

The research questions used throughout this study were:

- What are the motivations of airmen who attend the AF-TA briefing at FAFB?
- Are there significant demographic trends present for airmen who attend the AF-TA briefing at FAFB?
- What are the similarities between the motivations of airmen who attend PD courses at FAFB?

Participants

The specific demographics of FAFB are not publicly accessible; however, the USAF has an average age of 29 for enlisted personnel, has a population that is 70 percent White, and is male dominated with only 21.3 percent of the force represented by women.⁵² Moreover, roughly 90 percent of the enlisted population does not possess a bachelor's degree. It is possible that the demographics of the population at FAFB differ from the overarching demographics found within the USAF; however, FAFB is considered a standard operational base with all the major functional career fields represented and no clear evidence to support a drastic deviation from the overarching USAF demographics. The demographics of individuals who attend PD courses are currently unknown, and the collection of this data

⁵⁰ Blunt and Yang, "Factor Structure of the Adult Attitudes Toward Adult and Continuing Education Scale and Its Capacity to Predict Participation Behavior," 299–314; and Andrea O'Connor, "Reasons Nurses Participate in Continuing Education," *Nursing Research* 28, no. 6 (1979): 354–59.

⁵¹ *Personnel Assessment Program*, Department of the Air Force Manual 36-2664 (Washington, DC: Department of the Air Force, 2024), 58–66.

⁵² "Air Force Demographics," U.S. Air Force, accessed 18 February 2024.

during this study may spark additional questions or highlight unknown disparities.

Measures and Instruments

The EPS is an effective tool for identifying and measuring the construct of motivation for airmen who are pursuing PD at FAFB. The EPS has an extensive history measuring motivations of adult learners.⁵³ The scale has undergone two revisions before culminating in its final version of a 42-item questionnaire utilizing a four-point Likert scale.⁵⁴ The focus of questions range from “to improve language skills” to “to advance my mind.” The EPS is divided into seven subscales that assess potential motivations for adult learners: (a) communication improvement, (b) social contact, (c) education preparation, (d) professional advancement, (e) family together, (f) social stimulation, and (g) cognitive interest. The target population at EAFB consists of adult learners who are voluntarily pursuing education; however, there is no publicly accessible research that includes the use of the EPS on a military population.

Procedure

Formal USAF guidelines restrict the use of administering surveys to Department of Defense (DOD) personnel during their duty hours, through DOD equipment, on DOD property, and for the partial fulfillment of degree requirements outside of AU-sponsored programs.⁵⁵ The Air Force Survey Office conditionally approved this project with the following stipulations: (a) participants will complete the survey during their off-duty time, (b) the survey will be administered and completed outside of government facilities through non-USAF email accounts, (c) participants will be informed of their ability to participate in the study either during

⁵³ Blunt and Yang, “Factor Structure of the Adult Attitudes Toward Adult and Continuing Education Scale and Its Capacity to Predict Participation Behavior,” 299–314; Boshier, “Psychometric Properties of the Alternative Form of the Education Participation Scale,” 150–67; and O’Connor, “Reasons Nurses Participate in Continuing Education,” 354–59.

⁵⁴ Boshier, “Psychometric Properties of the Alternative Form of the Education Participation Scale,” 156–61.

⁵⁵ *Personnel Assessment Program*, 58–66.

a break or at the completion of the PD seminar, and (d) all requirements from the USAF Institutional Review Board (IRB) are met.

Approval and acknowledgement for all aspects of this project was formally obtained from the FAFB chief of education, the FAFB force support squadron commander, the FAFB legal office, the FAFB installation commander, the Air Force Survey Office, the Department of the Air Force Component Office of Human Research Protections, and the Johns Hopkins University Homewood IRB. This study was performed in accordance with the ethical standards expressed in the 1964 Declaration of Helsinki.

A purposeful sampling procedure was used, where all airmen who attended the AF-TA briefing at the FAFB education center during a two-month period were invited to participate in the study. The approximate number of airmen who attended the AF-TA briefing at the FAFB education center during this two-month period was 120 individuals. Airmen who attended the PD course were informed of the project after the completion of the seminar. Participants were informed of the study by a FAFB education center technician who read a preapproved script. Additionally, flyers with a QR code that linked to the online survey was available in the front of the classroom. Individuals who agreed to participate in the study and complete the survey had the opportunity to provide their contact information through a separate survey (with results that did not link to the primary survey) to receive a \$5 Amazon gift card.

Data Collection

The EPS was administered to participants through a survey link accessed through a QR code. Participants were asked to complete a demographics section that noted the military rank they hold, how long they have been assigned to FAFB, and their current education level, gender, and age. The results were collected online through Qualtrics. The participants were not asked for their name or any personally identifiable information. Downloaded files of the data were stored on a password-protected computer.

Results

Four airmen participated in the survey and completed all the questions. All participants were active-duty men and attended the seminar on different days. Two participants held the military grade of E-4 (a junior tier) and two participants held the military grades of E-5/E-6 (a supervisor tier). One participant's highest level of education was a high school diploma, two held associate degrees, and one held a bachelor's degree. The participant's length of time stationed at FAFB ranged from less than six months to between three-to-four years. Two participants identified as White, one identified as a native Hawaiian or other Pacific Islander, and one identified as other.

The seven subscales of the EPS were calculated for each of the participants by averaging their response to the respective six questions within each subscale (table 1). The lowest possible score is 1.00, indicating no influence of this motivational factor. The highest possible score is 4.00, indicating a significant influence of this motivational factor. No outliers were identified from within the data.

Additionally, the mean scores were identified across all participants and showcased "Professional advancement" as the overall highest indicated motivation with $M=3.50$ and "Social contact" as the overall lowest indicated motivation with $M=1.46$ (table 2). The primary, secondary, and tertiary education participation motivations were identified for each participant individually and as a collective based on the overall mean scores (table 3). In the event of a tie, both motivations with the same score were listed.

Further analysis of the results by demographics was not conducted based on the low volume of participants in this study. The demographic makeup of the participants represents a wide range of experience, time stationed at FAFB, and education level; however, this range paired with a low volume of participants limits the ability to identify meaningful correlations between demographic categories. Future studies with an expanded participant pool should be conducted to identify and analyze USAF voluntary education participation motivations trends between demographic categories. It is possible that an expanded study would reveal

Table 1. Student scores on the EPS subscales

	Student 1	Student 2	Student 3	Student 4
Communication improvement	2.83	1.67	1.00	2.00
Social contact	2.17	1.00	1.67	1.00
Educational preparation	2.33	2.33	2.50	2.83
Professional advancement	3.00	3.83	4.00	2.17
Family together	2.33	1.00	2.17	1.83
Social stimulation	2.00	1.17	1.67	2.50
Cognitive interest	3.17	2.17	2.17	2.67

Source: courtesy of the author.

Table 2. Mean and standard deviation on subscales in the EPS

	X scale		
	n	M	SD
Communication improvement	4	1.88	0.76
Social contact	4	1.46	0.57
Education preparation	4	2.50	0.24
Professional advancement	4	3.50	0.49
Family together	4	1.83	0.59
Social stimulation	4	1.83	0.56
Cognitive interest	4	2.54	0.48

N = number of students who answered each item.

Source: courtesy of the author.

correlations between demographic categories and voluntary education motivations; however, the analysis of a limited sample size of four participants can produce preliminary findings to shape the design and scope of future research studies.

Table 3. Primary, secondary, and tertiary education participation motivations by student

	Primary	Secondary	Tertiary
Student 1	Cognitive interest	Professional advancement	Communication improvement
Student 2	Professional advancement	Education preparation	Cognitive interest
Student 3	Professional advancement	Education preparation	Family together and cognitive interest
Student 4	Education preparation	Cognitive interest	Social stimulation
Overall	Professional advancement	Cognitive interest	Education preparation

Source: courtesy of the author.

Discussion

Even though the survey was administered during a PD seminar designed for education preparation, professional advancement had the highest mean score overall. This result could be an indication that airmen at FAFB are more motivated to pursue voluntary education because they believe it will help their career in the USAF, not necessarily because they are interested in the content offered through educational programs or as a steppingstone for future educational pursuits.

Despite a small sample size, six of the seven possible motivations to participate in voluntary education appeared within airmen’s top three motivations. Due to the small sample size, this finding cannot be applied to a broader population; however, the preliminary results from this limited study indicate a potential diversity in motivations to pursue voluntary education that should be explored in additional studies with a broader participation pool. Moreover, cognitive interest was the only motivation to be present in all respondent’s top three motivations. This variation highlights that airmen may be participating in the same USAF PD pro-

gram for different reasons and can have differing desired outcomes for participating in voluntary PD.

It is possible that if the PD seminar at FAFB included and advertised clear avenues for professional advancement, a greater number of airmen would use the program. The primary collective motivation of professional advancement, despite the course being designed for education preparation, may be an indication of a desired shift in program focus/outcomes from airmen. Future studies that include a control group should be conducted to analyze this possibility.

Recommendations

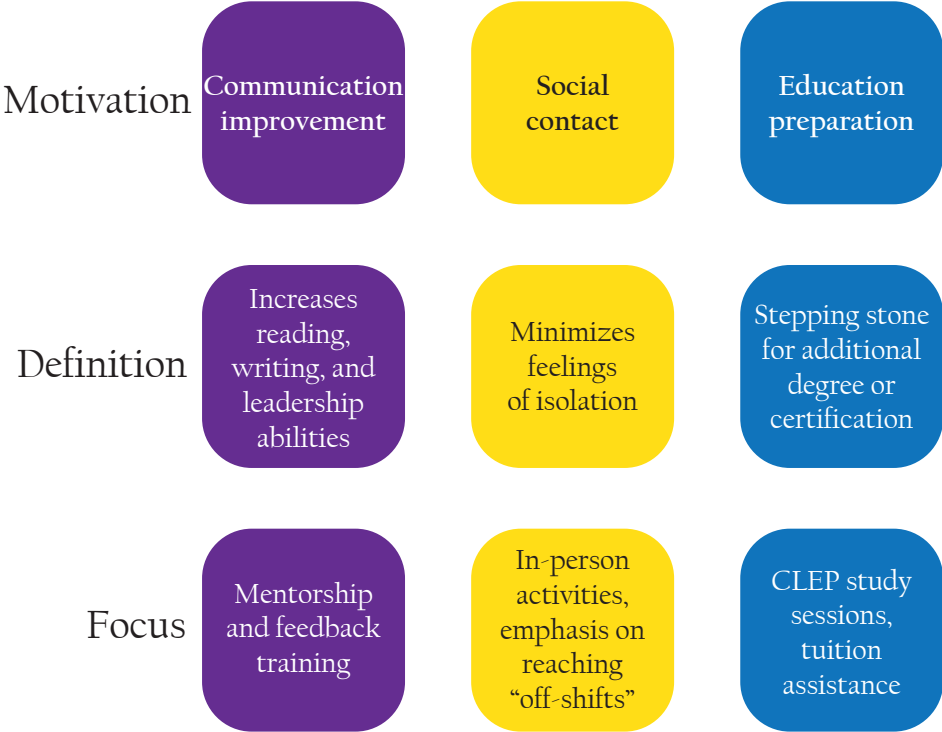
This study showcased the variety of motivations to pursue voluntary education present on a USAF installation and explored the multiple benefits of PD within the military. Despite the demonstrated success of military PD, airmen typically have limited options for pursuing PD on USAF installations. Additionally, the variability in motivations to pursue PD could drastically differ between military installations. Base leaders who want to optimize their PD programs will first need to understand the motivations of their specific population to pursue voluntary education.

It is possible for a base to offer PD programs that target each of the seven motivations to pursue voluntary education. Figure 1 outlines a model of seminars that a military PD program could incorporate to reach this goal. PD programs that align with the motivations to pursue voluntary education could already exist on a military installation throughout a variety of organizations, to include the education center, military and family readiness center, career assistance advisors, and private base organizations/clubs; however, they can belong to different reporting chains and operate exclusively of each other.⁵⁶ A wholistic view of PD on a military installation that includes all organizations providing a service to develop personnel could strengthen the overall PD effectiveness.

Military bases should encourage collaboration between their PD programs through a centralized organization. An overarching PD program that includes all organizations on an installation that provide PD could

⁵⁶ “Resources: Air Force Services Center,” MyAirForceLife.com, accessed 21 February 2024.

Figure 1. Military professional development model using the Education Participation Scale



Source: courtesy of the author, adapted by MCUP.

foster mutual support, use combined incentives, and streamline volunteer positions. Organizations that provide PD on a military base could deconflict their scheduled times for seminars to ensure that servicemembers can attend as many programs as possible. Additionally, opportunities for PD on base could be advertised during each seminar regardless of the specific organization that is hosting the course. A centralized attendance tracker between all PD seminars could be utilized to offer incentives for participating in PD (e.g., a servicemember receives a certificate after 10 courses and a challenge coin after 15 courses). Lastly, a combined PD program could streamline manpower by centralizing advertising, curriculum development, and instructor training. The centralized program leaders could analyze the PD course offerings across the installation, assess the

Professional
advancement

Family
together

Social
stimulation

Cognitive
interest

Target skills
beneficial
for promotion

Incorporates
seminars
useful to
military
dependents

Break
from the
day-to-day
job

Subject
material that
appeals to
personnel

Performance
reports or
award
packages

New parent,
emphasis
on fun events

Increased
frequency
of PD
options

Unique
courses
with subject
matter
experts

ability of the base to offer courses that match the motivations of their population to pursue education, and advocate for additional resources or seminars, as needed.

Conclusion

The U.S. military has a rich history of using PD for a variety of purposes. PD is beneficial to both the institution and individual servicemembers.⁵⁷

⁵⁷ Classen, Horn, and Palmer, “Needs of Military Families,” 250; Kim and Won-kwang, “The Effects of Joint Education Programs on Cohesion among South Korean Military Academy Cadets,” 517–18; and Jordan and Curtis, “Evaluating the Impacts of Technology Education on Military Maintenance Students,” 41–42.

Effective PD requires clear institutional support and a perception of quality.⁵⁸ Institutional barriers, such as a negative perception of military PD, could be mitigated through a deliberate focus on matching PD programs to the motivations of the local base population to pursue voluntary education.⁵⁹ This study demonstrated a variety of motivations to pursue voluntary education at one USAF installation. Based on the results, it is recommended to leverage the focus of multiple PD programs on base into a centralized approach. A unified PD program could be used to deconflict schedules, advertise courses, create combined participation incentives, and streamline limited volunteer manpower. Additional studies should be conducted to analyze the motivations of servicemembers to pursue PD at additional military installations, the differences in motivations to pursue PD between military installations, and the effectiveness of a centralized PD program.

⁵⁸ Makopoulou et al., “An Investigation into the Effects of Short-Course Professional Development on Teachers’ and Teaching Assistants’ Self-Efficacy,” 790–91; and Newcomer et al., “Just Checking the Box,” 30–45.

⁵⁹ Boshier, “Psychometric Properties of the Alternative Form of the Education Participation Scale.” 150.



CHAPTER 7 ●

Making the Case against Unimodal Instructional Strategies

A Qualitative Intrinsic Case Study
within Instructor Upgrade Training

Major Brad Kelly (USAF), EdD

At the 128th Air Control Squadron, a control and reporting center (CRC) within the Wisconsin Air National Guard, air battle manager (ABM) instructors struggle with the fundamentals of serving in a teaching role. Most notably, during the unit's most recent deployments in support of operations in the skies over Iraq, Syria, and Afghanistan, a common deficiency in instructor proficiency was noted between the Air National Guard (ANG) and their active duty counterparts while deployed in a Joint operational environment. This observation led to the implementation of this study and examination of all possibilities of a deficiency in either overall training standards or requirements for the ANG's instructor upgrade (IUP) training program.

The focus of this study was to identify the most significant areas of struggle that ANG ABM instructors experience and, more importantly, gain a better understanding of their perceptions toward the overall effectiveness of the current IUP training program. To uncover areas for growth in the current training, the researcher utilized a qualitative intrinsic case study in a three-phased approach of analyzing artifacts and conducting both observations and semistructured interviews. The researcher rooted this study methodology and design in a theoretical framework most closely associated with social constructivism and andragogical principles of adult learning.

The researcher of this study has carefully vetted all information contained within this research to ensure that no information designated as for official use only or controlled unclassified information is reported. Specific numbers and training course names, tactics, techniques, and procedures related to the ABM career field and training program are rough estimates and do not disclose the exact training undergone by U.S. Air Force ABMs. The findings of this study indicate a lack of proficiency due to both an overreliance on unimodal instructional techniques and a lack of application of andragogical principles in the current training syllabus. Additionally, the research highlighted academic content areas that are currently nonexistent and contribute to ABM control ability deficiencies. The remainder of this chapter details the methodology of the study to pass along best practices in research design that may be applicable to other Service branches, specific career fields, and training environments within the U.S. military.

Methodology

Introduction: Research Questions

This section details the methodology of the study and the researcher's use of a qualitative intrinsic case study design to uncover the perceptions of ABM instructors on the use of unimodal instructional techniques and the related deficiencies in the current ABM instructor upgrade training program within the ANG. To clarify, ABMs are Air Force officers who use strategy and intimate knowledge of aircraft weapons and surveillance capabilities to control the outcome of an aerial battle. They are responsible for prioritizing aircraft weapons, sensors, and fuel within the acceptable level of mission risk to achieve the respective commander's intent for the air battle. They can perform this role from either an airborne or ground surveillance control station and are not to be confused as pilots, a completely other career field within the Air Force. The role of an instructor within the ABM career field is to serve as the unit's subject matter expert (SME), lead educator, and tactician in the art of tactical aircraft control. Last, this study explicitly researched the IUP training program among ABMs within the ANG, a reserve component or branch within the broader Air Force.

Additionally, the purpose of this research and methods section is to examine and conduct qualitative analysis within the researcher's ANG squadron and ABM career field. The qualitative data collection and analysis in this chapter focus on the topics covered within the literature review section of the study consisting of an overreliance on a singular instruction method such as Microsoft PowerPoint, andragogical principles for military learners, and the benefits of implementing differentiated instruction techniques in current military training programs.

Within this study, the researcher aimed to understand the current training program, where ABM instructors struggle and feel unprepared by the current set of training requirements. The findings of this research describe what methods of instruction can be improved to increase the overall quality of ABM instructors' education. The researcher also examined what perceptions toward their training exist among ABM instructors that might affect instructor proficiency. Last, in addition to identifying and analyzing what strengths and deficiencies in training currently exist, the researcher also determined how the ANG's IUP training program standards could align with the active duty component requirements. This alignment between the two components' training programs is vital for congruence and consistency of training, ensuring that members from each component are trained similarly and have achieved the same proficiency level when deployed in a Joint operational environment. Therefore, the researcher established the following three research questions to provide the foundation necessary to conduct this study:

- What strengths and weaknesses currently exist in the ANG's IUP training program due to general program requirements, current training methods, and overall program structure?
- What perceptions exist among ABM instructors toward the current training program with respect to the use of unimodal instruction and andragogical principles within their training?
- What changes in overall program structure, instructional delivery, and/or content are necessary to promote

improved instructor upgrade training for ABM to better align the ANG with the active duty?

Theoretical Framework Application

Malcolm Knowles' principles of andragogy and the theory of social constructivism directly related to and informed the research questions of this study. Specifically, the researcher used these two theoretical frameworks as they set out to determine if the current ABM IUP program allows for choice in learning through differentiated instruction or if a unimodal instructional technique is effective in training ABM instructors. In line with the first theory, Knowles suggested that adult learners are vastly different from children in terms of self-motivation and overall knowledge attainment.¹ For example, adults need learning to be directly beneficial and related to the skills required to be successful in their current professions. Adults are also motivated by using these skills to directly benefit and help those around them, compared to extrinsic motivations, such as awards, which are often preferred by children. Furthermore, and in line with the second theory the researcher used, according to John W. Creswell and Cheryl N. Poth, "social constructivist frameworks act for societal improvements through co-created findings between the researcher and participants within an organization."² The theory of social constructivism applied to this study because the findings directly affected and focused on improving the local air control squadron training processes in which the researcher worked.

The first research question analyzed the current ANG IUP training program to determine deficiencies in program requirements and training methods. This was followed by the second research question, which analyzed the perceptions of the ABM learners toward the current program and choice of training methods. Last, the theory of social constructivism in benefiting others in future programs informed the third research ques-

¹ Malcolm Knowles, *Andragogy in Action: Applying Modern Principles of Adult Learning* (San Francisco, CA: Jossey-Bass, 1985).

² John W. Creswell and Cheryl N. Poth, *Qualitative Inquiry and Research Design: Choosing Among Five Approaches* (Thousand Oaks, CA: Sage, 2018), 34.

tion related to closing the gap between the ANG and active duty instructor upgrade training programs.

Additionally, the researcher applied both constructivism and the theory of andragogy in collecting and analyzing the data throughout this study. *Constructivism* refers to the “approach to learning that describes that people actively construct or make their own knowledge and that reality is determined by the learner’s experiences.”³ Additionally, according to leading experts in the field of constructivism, the central idea of constructivism is that “human learning is constructed and that learners build new knowledge upon the foundation of prior learning. This prior knowledge influences what new concepts an individual will construct from the new learning experiences.”⁴ This building on or constructing new learning concepts from past experiences is especially evident in military training. For example, much of the military training environment begins with teaching basic skills and building on or constructing more elaborate skills and knowledge as one progresses in their career. Therefore, the researcher used constructivism theory in data collection and analysis to look for specific examples where the current training program is building on prior knowledge and continually evolving to better the training of ABMs.

Furthermore, the assumption is that by the time an ABM reaches the point in their career to become an instructor, they have already acquired the prior skills and knowledge necessary to become a SME in the field, in line with constructivist principles. Therefore, the research questions of this study addressed potential shortfalls where ABMs did not possess prior knowledge, and the perceptions of ABMs related to current instructional delivery methods. In determining these shortfalls or gaps in knowledge, prior skills of instructors were highlighted during data collection using observations and semistructured interviews to determine the previous attainment of knowledge in specific areas. In analyzing data that reflected where deficiencies in ABM instructor knowledge, the researcher

³ Stephen N. Elliott, Thomas R. Kratochwill, Joan Littlefield Cook, and John F. Travers, *Educational Psychology: Effective Teaching, Effective Learning* (Boston, MA: McGraw-Hill, 2000), 172.

⁴ Denis Phillips, “The Good, the Bad, and the Ugly: The Many Faces of Constructivism,” *Educational Researcher* 24, no. 7 (October 1995): 9, <https://doi.org/10.2307/1177059>.

utilized the emerging themes to inform specific areas that the ANG could use to improve the training of ABM instructors.

The researcher also applied the theory of andragogy to this research, in which the ideas and themes centered around adult military learners became the focus of the research questions, data collection, and data analysis. The researcher used the andragogical principles of Knowles to frame the associated research questions examining whether the ANG applied the best academic training practices toward adult military learners. Specifically, Knowles described andragogy and adult learners as preferring an active choice in how they learn.⁵ Therefore, using semistructured interviews, the researcher examined the perceptions of ABM instructors toward the use of unimodal instruction in the current training program. The researcher also explored the current gaps in knowledge among ABM instructors and examined if the gaps were due to the current mode of instruction. The application of the research questions directly aligns with Knowles's principle of andragogy, in which adults prefer their learning to have an immediate and direct impact on their daily tasks. Any gaps in knowledge discovered during the research will inform the current IUP program by identifying areas where the program can be supplemented and better aligned to the more robust training of other careers in Air Force career fields.

In summary, the two theories of constructivism and andragogy informed the data collection and analysis by providing a lens to determine and examine emerging themes. The data collection and analysis ultimately related to the purpose of this study by identifying the deficiencies of the current training program, perceptions of the ABM instructors toward that program, and informing changes to the program to improve future instruction.

Research Design and Rationale

The research conducted during this study was a qualitative intrinsic case study design. According to John W. Creswell, qualitative research allows

⁵ Svein Loeng, "Various Ways of Understanding the Concept of Andragogy," *Cogent Education* 5, no. 1 (January 2018): 1–15, <https://doi.org/10.1080/2331186x.2018.1496643>.

the researcher to explore or seek a deeper, more detailed understanding of a complex phenomenon or topic of study.⁶ Robert K. Yin indicates that case studies are a qualitative inquiry design in which the researcher constructs a thorough analysis of a case, often in the form of evaluating a program, event, or one or more individuals.⁷ In this specific research, the case the researcher analyzed was the current instructor upgrade training program used within the ANG at the 128th Air Control Squadron.

Furthermore, this study is an intrinsic case study. According to Creswell and Poth, “intrinsic case studies focus on the case itself (i.e., evaluating a program) because the case presents an unusual or unique situation.”⁸ The overarching case in this research is the IUP program. According to Robert E. Stake, researchers of case studies collect information through multiple data collection procedures.⁹ While conducting this qualitative research, the researcher collected data using multiple procedures, such as analyzing prior instructor performance records and other artifacts, conducting observations, and leading semistructured interviews with the participants.

The multiple forms of data allowed the researcher to examine the participants’ shared experiences and uncover any existing deficiencies in the ANG’s IUP training program. The researcher collected qualitative data through semistructured interviews to determine the participants’ thoughts regarding why they felt either prepared or unprepared as instructors and what improvements were needed to train them better in their role as instructors and eliminate any deficiencies in the current instructor upgrade training program. Additionally, the researcher examined what factors potentially contributed to these deficiencies within the current training program and the perceptions of ABM instructors toward the current training methods. The researcher then discussed recommendations and the need for specific supplemental academic topics to the cur-

⁶ John W. Creswell, *Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research* (New York: Pearson Education, 2005).

⁷ Robert K. Yin, *Case Study Research: Design and Methods*, 3d ed. (Thousand Oaks, CA: Sage, 2009).

⁸ Creswell and Poth, *Qualitative Inquiry and Research Design*.

⁹ Robert E. Stake, *The Art of Case Study Research* (Thousand Oaks, CA: Sage, 1995).

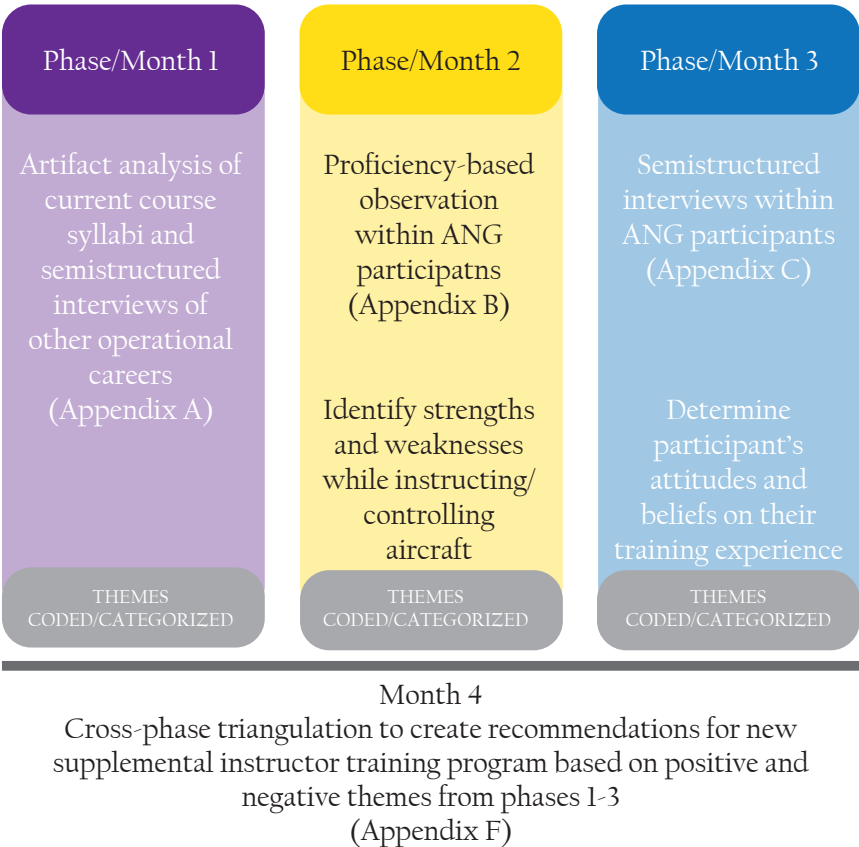
rent training program to correct these deficiencies and avoid noticeable differences between the competency levels of ANG instructors and their active duty counterparts. These deficiencies and differences in instructor ability were first noticed during one of the unit's most recent deployments, and this research aimed to gather information and data about the current training program, which may be leading to these deficiencies.

This qualitative intrinsic case study research design allowed the researcher to dig deeper into the perceptions about the current training program by probing the thoughts, feelings, and beliefs of those who recently completed the IUP training program. Current instructor opinions on program effectiveness ultimately impacted recommendations for a new supplemental training program. Similarly, the researcher collected data on the attitudes and beliefs of current instructors as to what areas or deficiencies exist that led to them feeling unprepared to conduct their job as instructors. As a result, the researcher felt it imperative to use a qualitative design because it is essential to understand the participants' attitudes and beliefs about their previous learning experiences before proposing significant changes in current training processes.

The first phase in collecting data for this study (figure 1) began with identifying multiple artifacts, specifically the ANG and active duty course syllabi, and conducting semistructured interviews to create a baseline understanding of the training programs between the two Air Force components and other operational career fields. The researcher utilized a semistructured interview protocol to conduct all interviews during phase one of the research with members of various operational Air Force careers. In this first phase of the research, the intent was to determine similarities and differences in program requirements, instructional methods, and overall academics between the various IUP within the Air Force.

During phase two of the study, the researcher then examined the ANG participants' current proficiency level by observing their air control missions as they serve in an instructor role to other air traffic controllers. The researcher conducted observations using a standard Air Force "Individual Mission Gradesheet" in Air Combat Command (ACC) Form 206 to mitigate any subjectivity in the observational phase and help determine student proficiency levels. Air battle manager instructors use Form 206

Figure 1. Research design diagram



Note: the researcher constructed this figure to depict and aid in visualizing the different phases of the study and overall qualitative research design.
Source: courtesy of the author, adapted by MCUP.

any time a control mission requires an instructor to be present during the mission. These observations also helped identify potential areas of improvement in participant performance, indicating possible shortfalls in the current instructor training curriculum. The researcher then coded and categorized the information to develop themes for areas of improvement. However, areas that were performed correctly needed to remain in the current training program and were also noted and considered in developing the new supplemental training program.

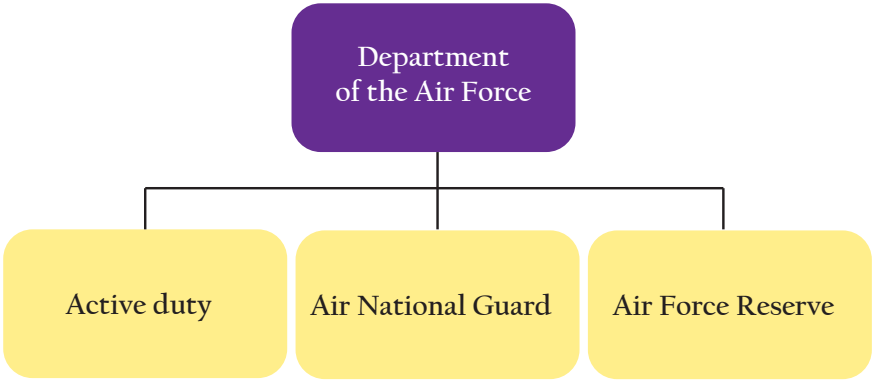
Once these observations were complete, the researcher utilized semistructured interviews in the third phase of this study to share the results with the participants and ask a series of questions relating to their performance. The researcher also asked participants what areas they felt needed improvement or needed to be maintained in the current training program. The researcher then coded, categorized, and finally themed the results from the semistructured interviews to analyze this qualitative data. Last, once the researcher identified recommendations based on the feedback consisting of the qualitative data and areas of either improvement or strengths, the researcher provided recommendations for a new supplemental instructor training program curriculum to close gaps in current deficiencies. This curriculum will better align with active duty tactical aircraft controllers and implement elements of andragogy and differentiated instruction to best train adult military ABMs within the ANG.

Site Selection and Participant Sampling

This study was bound by location, analyzing current ABM training within the ANG and 128th Air Control Squadron. To best understand the research and analysis of this topic, an introductory lesson in specific career field organizational structures and critical Air Force language is necessary. The Air Force is organized into three subcomponents (figure 2). The participants of this research serve in both the operational career fields of the active duty component and the ANG component, a reserve component of the broader Air Force. The ANG branch is not to be confused with the Air Force Reserve, which is an entirely separate third branch of the Air Force.

Furthermore, the participants of this study are all members of what the Air Force calls operational career fields, which indicates direct involvement in flight operations. Specifically, an ABM's sole job is to prioritize aircraft weapons employment, aerial surveillance sensors, and fuel distribution for the operational aircraft under their control. These ABMs, who work at air control squadrons, are in constant contact with military aircraft via radios in the pilot's cockpit and visually detect aircraft using the ABM's tracking radar.

Figure 2. Overarching structure of the U.S. Department of the Air Force



Source: courtesy of the author, adapted by MCUP.

The researcher analyzed the quality of the ABM instructor upgrade training programs within the two previously mentioned Air Force components. To further specify, an ABM advances in their career by eventually becoming a certified instructor once they have completed the mandatory IUP training. The air control squadrons view these instructors as the tactical experts within the organization, and after becoming certified, they are relied on to teach other ABMs academics and evaluate mission performance. Overall, ABMs are responsible for ensuring proper and efficient management of the aerial battle picture, and more importantly, the pilots under their control remain safe and return home to their loved ones.

Site

This study took place at the 128th Air Control Squadron, located at Volk Field Air National Guard Base at Camp Douglas, Wisconsin. The researcher chose this site due to a vested interest in improving the current training and educational processes at the researcher's place of employment. Additionally, although nine total ANG air control squadrons exist across the United States, the 128th Air Control Squadron is the only one within the researcher's state of residence or surrounding states. The limited numbers of other air control squadrons within proximity of one another present

Table 1. Air National Guard air control squadrons

Air Control Squadron (ACS)	Location
103d ACS	Orange, CT
109th ACS	Salt Lake City, UT
116th ACS	Warrenton, OR
117th ACS	Savannah, GA
123d ACS	Blue Ash, OH
128th ACS	Camp Douglas, WI
134th ACS	Wichita, KS
141st ACS	Aguadilla, PR
255th ACS	Gulfport, MS

Source: courtesy of the author, based on data from “Control and Reporting Centers (CRC),” fact sheet, Air Combat Command, 20 March 2023.

added difficulty for research at multiple sites. Table 1 shows the various ANG air control squadrons and their respective locations.

The 128th Air Control Squadron’s mission statement is “Command, control, communication, computer, and intelligence professionals providing aerospace battle management for any tasking, anytime, anywhere.” Additionally, the 128th Air Control Squadron’s vision statement is as follows, “Embracing the 21st-century as a Battle Management Team Focused on Excellence, Vision, and Integrity.” Considering this organization’s mission and vision statement, the researcher conducted this study at this study site to ensure the organization was training ABMs to be world-wide deployable at a moment’s notice and had the most current twenty-first-century training and skills necessary to succeed in combat.

The 128th Air Control Squadron’s tactical operations occur on an open-layout floor, with multiple computer consoles designed to track military aircraft as training in tactical aerial battle missions is conducted. Further, other nonoperational staff perform various functions, from maintenance support to more clerical and administrative support to all military members of the unit. Additionally, the 128th Air Control Squadron is structured administratively similar to all other ANG air control squad-

rons regarding staffing numbers, positions, and roles and responsibilities, as determined by higher headquarters. For example, pending current recruiting and retention numbers, the 128th Air Control Squadron comprises approximately 35 full-time staff members who work Monday through Friday every week, and the remaining 150 National Guard members of the unit are considered one weekend a month drill status guardsmen. The unit consists of three separate primary sections: operations, maintenance, and the command staff. This study focused on the operations section, consisting of approximately 60 members, and the operations section members classified and trained as ABMs, with an Air Force specialty code (AFSC) of 13B. Of the approximately 60 members within the operations section, only 7 are trained as ABMs and go through the instructor upgrade training program, which was the focus of this study.

Participants

The sampling strategy used to conduct this research was purposeful and considered convenience sampling, as all participants were required to be current Air Force officers in various operational career fields. For example, participants included active duty officers from various career fields and qualified ABM instructors at the 128th Air Control Squadron. The benefit of using this strategy was that the participants were quickly and easily identifiable, and due to the participants being coworkers of the researcher, ease of access to the participants was attainable. According to Creswell and Poth, using purposeful sampling is beneficial due to being able to intentionally select multiple participants to describe an event, process, or program, such as in this study.¹⁰ Additionally, the benefits of using convenience sampling are that it is cost-effective, efficient, and simple to implement.¹¹ However, using this sampling strategy and the researcher knowing all participants personally opens the door for bias in

¹⁰ Creswell and Poth, *Qualitative Inquiry and Research Design*.

¹¹ Justin Jager, Diane Putnick, and Marc Bornstein, "More than Just Convenient: The Scientific Merits of Homogeneous Convenience Samples," *Monographs of the Society for Research in Child Development* 82, no. 2 (May 2017): 13–30, <https://doi.org/10.1111/mono.12296>.

research, particularly in a military environment where the rank structure and chain of command can also influence responses and openness from the sampled participants.

This study consisted of two groups of participants, which the researcher used in different phases of this research. The first group of participants in the first phase of this research consisted of five individuals from the active duty component of the Air Force and of varying operational career fields. The first group of participants had five different AFSCs and had previously attended IUP training within their respective functional areas. The group consisted of three different pilots of varying aircraft platforms, one intercontinental ballistic missile operator, and, most importantly, one active duty ABM instructor. The researcher chose these active duty individuals via convenience sampling as they were all chief of weapons and tactics counterparts of the researcher within their respective units and careers. This group of individuals was contacted and used in the first phase of data collection to determine the instructor upgrade training programs and academic standards for other operational career fields and, most importantly, among the active duty ABM career field. Last, the researcher interviewed these five participants in the first phase of the study using the semistructured interview questions to understand the baseline IUP requirements among active duty members.

The researcher also chose the second group of participants using convenience sampling, who were among the researcher's current coworkers at the 128th Air Control Squadron. The second group consisted of only ANG ABM participants, which was made up of full-time staff instructors or department heads within the organization and the instructors who are traditional drill status guardsmen, who only report to their military job one weekend each month. The full-time staff or department heads consisted of two individuals, and the part-time staff consisted of another two individuals, for a combined total of four participants. Reference table 2 below for a breakdown of participants based on Air Force career field, service component, years in profession, and years spent instructing in their respective career field.

Table 2. Participant demographics

Participant # /study phase	AFSC	USAF component	Years in profession	Years instructing
1 – Phase 1	Pilot – 11F	Active duty	7	2
2 – Phase 1	Pilot – 11T	Active duty	8	3
3 – Phase 1	Pilot – 11A	Active duty	6	2
4 – Phase 1	ICBM – 13N	Active duty	5	1
5 – Phase 1	ABM – 13B	Active duty	9	4
6 – Phase 2 & 3	ABM – 13B	Air National Guard	6	1
7 – Phase 2 & 3	ABM – 13B	Air National Guard	8	3
8 – Phase 2 & 3	ABM – 13B	Air National Guard	4	2
9 – Phase 2 & 3	ABM – 13B	Air National Guard	7	4

Source: courtesy of the author.

Data Collection Procedures

The data collection process of this study consisted of three phases, each requiring one month to complete and occurring during the respective month’s drill period weekend for the ANG unit. Prior to data collection, the researcher acquired all permissions within the study site from both the organizational leadership and study participants. Current leadership within the study site granted permission to conduct this research and were eager to see the final results of this study to determine what improvements were necessary and whether additional training changes would benefit the current employees of the 128th Air Control Squadron. As a result, the current organizational leadership was entirely supportive in all research steps. Furthermore, all participants signed informed consent forms agreeing to contribute to this research. Similar to the enthusiasm expressed by the organizational leadership, all ANG participants were eager to participate. Due to their vested interest in the organization, they were also anxious to see this study’s final results, which would improve their organization.

Research Phase 1:

Artifact Analysis and Active Duty Semistructured Interviews

In the first phase of data collection, the researcher examined the current ABM instructor upgrade training program within the ANG by using artifacts written in the form of governing training regulations, current course syllabi, and academic training material in the form of PowerPoint slides. The researcher utilized artifact analysis within this study due to the benefits of not requiring human participants, which can prevent unexpected results, and the researcher can use artifact analysis to study people across both time and space.¹² Using the artifacts, the researcher identified what methods of instruction the ANG is currently implementing (i.e., PowerPoint briefs, assigned supplemental readings, and short instructional videos). The researcher then examined the current IUP syllabus to determine overall requirements for how many hours of both instruction and academics were required to qualify an ANG ABM as an instructor. The researcher also focused on what prerequisites were currently required to determine admission into the instructor upgrade training program.

In the second half of phase one, the researcher then compared these details on the current instructor training program within the ANG to the active duty ABM training program and all other IUP training programs in other operational career fields to determine any similarities or differences. To discover similarities and differences in the training among related career fields within the Air Force, the researcher conducted semistructured interviews with five different participants from the active duty component. All interviews with participants during phase one took place in a private office setting to ensure confidentiality.

Research Phase 2:

Observations

During the second phase of this study, the researcher observed their second group of participants consisting of only ANG instructors from the

¹² Bella Martin and Bruce Hanington, *Universal Methods of Design: 100 Ways to Research Complex Problems, Develop Innovative Ideas, and Design Effective Solutions* (Beverly, MA: Rockport Publishers, 2012).

128th Air Control Squadron. The researcher used live and simulated aircraft control missions to determine the current proficiency level and the knowledge base of ABM instructors at the 128th Air Control Squadron. The focus of this second phase was to determine what strengths and weaknesses the current instructors possessed when both controlling missions and instructing other less experienced ABMs. The researcher used an Form 206 to conduct these observations. Additionally, the researcher used mission observations to determine what areas instructors lacked proficiency in (i.e., proper intervention strategies, proper grading methods, various tactical aircraft control techniques, etc.) while performing or instructing tactical aircraft control missions. The researcher used the current Form 206 observation form for mission evaluation and used the form to create standardization and aid in the data collection throughout this study's observation phase. This observation protocol aided in the standardization of data collection and removed subjectivity in observational areas, which will help other researchers build on this initial research.

Research Phase 3:

Air National Guard Semistructured Interviews

The researcher then used one-on-one semistructured oral interviews in the third step of the study to understand the participants' own beliefs as to whether the IUP training program was preparing instructors for effective job performance and to understand what current instructors think was lacking in the instructor curriculum. According to Eric Drever, semistructured interviews are beneficial due to providing flexibility when used for small-scale research, such as this study.¹³ Additionally, similar to the phase one semistructured interviews, all interviews with participants took place in a private office setting to ensure confidentiality. As participants answered each of the protocol questions, the researcher took hand-written notes of their responses. Like phase one interviews, the researcher ended each interview by asking each participant if the notes

¹³ Eric Drever, "Using Semi-Structured Interviews in Small-Scale Research: A Teacher's Guide," *Scottish Educational Review* 28, no. 1 (January 1996): 81-82, <https://doi.org/10.1163/27730840-02801014>.

the researcher collected accurately reflected their respective feedback on each question. This member checking method increases validity due to the researcher ensuring they did not misinterpret the data that the participants provided by bringing the data analysis back to the participants for verification.¹⁴ The researcher then coded the participants' responses in the semistructured interviews to determine major themes related to strengths and areas that needed improvement in the current training program. According to Creswell and Poth, coding and placing data into various themes is central to qualitative research and involves making sense of the feedback collected during observations and interviews.¹⁵

Data Analysis Procedures

The researcher used a five-phase process to analyze the data. These phases included: prepare the data for analysis; explore the data; analyze the data; represent the data analysis; and interpret the results. Creswell and Poth urge the use of a sequential, step-by-step data analysis procedure and protocol for conducting qualitative research to segment the collected data.¹⁶ This segmentation and chunking in the analysis phase allow the researcher to more effectively manage and organize the plethora of data collected in a qualitative study to understand the findings.¹⁷ The researcher's use of a step-by-step and structured data analysis process ensured that all data was methodically taken into account and analyzed to determine the most accurate findings possible. Additionally, breaking down the data analysis phase into multiple steps enabled the researcher to stay organized and stay on a timeline, completing each step by a specific phase in the study.

Prepare the Data for Analysis

The researcher viewed the analysis of the collected data through a social constructivist, constructivist learning theory, and andragogical theoretical perspective. The specific data the researcher examined consisted of

¹⁴ Valerie Otero and Danielle Harlow, "Getting Started in Qualitative Physics Education Research," *American Association of Physics Teachers* 2, no. 1 (2009): 2.

¹⁵ Creswell and Poth, *Qualitative Inquiry and Research Design*.

¹⁶ Creswell and Poth, *Qualitative Inquiry and Research Design*.

¹⁷ Creswell, *Educational Research*, 150.

artifact data and semistructured interviews of active duty members from phase one. Then, in phases two and three, the researcher examined observational performance feedback and program opinions from semistructured interviews among the four ANG instructors. When analyzing the current instructor upgrade training program, the researcher emphasized the training program's overall structure and requirements, academics, and instructional methods. For example, the researcher used an *a priori* method to focus on how effective the current academics, teaching methods, and overall program structure are in relation to ABM instructor performance and interview feedback.

Initial notes were annotated via pen and paper on the course syllabi from phase one and the Form 206 observation forms from phase two. Additionally, the researcher annotated notes on the semistructured interview protocol. The researcher organized all data and associated notes on the respective phase protocol worksheets found in the appendices in preparation for analysis. Once complete with filling out all three of the phase protocol worksheets, the researcher stored this information in their personal locked safe in their office to ensure confidentiality of findings and maintain participants' privacy prior to beginning the data exploration.

Explore the Data

Once all artifact data, mission observation forms, and interview protocols were complete, the researcher then organized data by theme or commonalities and examined patterns in the participants' responses. For example, if a participant had feedback regarding the instructional delivery methods currently used, the researcher then placed that individual's feedback comments with other participants who had similar opinions or experiences regarding instructional delivery methods. To do this, the researcher used an Excel spreadsheet to record different commonalities and themes using the program's different worksheet tabs. Using this method, the researcher easily grouped participant observations and feedback. For example, observational trends collected throughout the mission or performance-based observational phase were also themed and categorized together among the participants. If a participant performed below

the standards outlined in the Form 206 in a category such as aerial refueling ability, the researcher placed their specific results with others who also performed below expectations in that category.

Conducting this hand coding based on observational strengths and weaknesses allowed the researcher to determine various insights into the course academics when cross-referenced with the data the researcher collected in phase one. According to Creswell, “the process of coding is central and at the heart of qualitative research and involves making sense of the text collected from interviews, observations, and documents.”¹⁸ The researcher also explicitly coded all their own notes throughout all phases of the research and used open coding to organize data into themes from all the ANG instructor participants to make sense of the data the researcher collected. Open coding is most beneficial when the researcher forms major categories of information about a phenomenon or situation and segments the information to help organize and manage the data.¹⁹ Additionally, placing data into themes allows the researcher to analyze the data more efficiently, especially for those early in their research career.²⁰ Specific to this study, these thematic areas helped identify strong components in the training program that were to remain in place or areas in the current training program that needed supplementation with more robust academics and an all-new supplemental training curriculum.

Analyze the Data

The first step in analyzing the data occurred by cross-referencing all phase one information pertaining to course syllabi and program requirements amongst the ANG, active duty, and other non-ABM operational career fields. To cross-reference the data within phase one, the researcher examined all codes that aligned to either similarities or points of difference between the ANG and active duty course syllabi. The researcher conducted this cross-referencing and artifact analysis to determine a baseline for

¹⁸ Creswell, *Educational Research*, 190.

¹⁹ Creswell and Poth, *Qualitative Inquiry and Research Design*.

²⁰ Virginia Braun and Victoria Clarke, “Using Thematic Analysis in Psychology,” *Qualitative Research in Psychology* 3, no. 2 (January 2006): 77–101, <https://doi.org/10.1191/1478088706qp0630a>.

how the ANG's program compared relative to other standards set forth by their active duty and other operational career field peers. The specific data the researcher analyzed explored program prerequisites prior to program entry, program requirements after entry of the student, specific academic topics, and the overall program venue and completion timeline.

In phase two of data analysis, the researcher used the Form 206s filled out during the observational phase and annotated specific performance strengths and weaknesses. For example, the researcher annotated strengths specific to tactical aircraft control fundamentals and certain weaknesses, such as a lack of proper student intervention strategies the ABM instructors used with their students. To annotate the specific strengths and weaknesses in performance, the researcher again used coding to break out each specific area of strength or weakness in control ability. Additionally, to verify that areas of weakness were consistent with instructors' knowledge level and not purely a performance mistake during the observation, the researcher used past performance evaluations to cross-check this information and determine if there was a history of a knowledge gap in these areas. The researcher obtained past Form 206 information for each participant, as this is considered open knowledge among already certified instructors and full-time members of the operations section within the squadron. Therefore, the researcher's employment status within the study site allowed for ease of access to these records.

As previously mentioned, the researcher annotated and grouped each strength and weakness in individual Excel workbooks created to reflect the specific participant, general observation, and respective strengths or weaknesses of the performance. Once all observations were extracted from the Form 206s and entered into the spreadsheets, each observation was coded based on performance category. For example, codes were assigned and consisted of performance during and with the initial aircraft check-in phase of the mission, aiding tactical fighter aircraft to receive aerial refueling from an aerial refueling-capable aircraft, such as the Boeing KC-135 Stratotanker, and other specific phases of the overall mission, such as the intercept phase and egress phase of the aerial fight. The researcher used each of these categories, which the Form 206 identifies and breaks down as the areas instructors must perform while instructing students.

Using this data analysis process of organizing the collected data by strengths, weaknesses, and specific themes related to aircraft control tactics, the researcher took account of all responses and comments received during the one-on-one semistructured interviews as well. First, the researcher entered participants' comments collected during the one-on-one interviews by placing handwritten notes into another Excel spreadsheet. The researcher annotated the participant's name, general comments, strengths or weaknesses, and recommendations. Once entered, the researcher grouped each participant's comments and feedback on the current ANG IUP training program with similar comments from other participants.

Represent the Data Analysis

The researcher used a narrative to represent the findings of the data analysis that the researcher conducted throughout this study. Once categorized, all major themes were detailed through each participant's words related to specific training strengths or weaknesses and in conjunction with the artifact analysis and participant observations from the other two phases of research. This narrative method the researcher utilized provides military leaders and the organization study site with a quickly identifiable means to understand each participant's personal feelings and critical recommendations for improving the current training processes.

Interpret the Results

Once each of the three phase-based spreadsheets or databases was complete, the researcher cross-checked each response and used triangulation between the three phases of research to discover any significant common themes and interpret the significant findings of this study. For instance, common themes included aircraft control academic topic areas, overall program standards and requirements, instructional delivery methods, and the use of andragogical principles within the program. Overall, the researcher interpreted the results found within each of the three research phases to decide which significant strengths in the current training program needed to remain in place. The researcher also interpreted the data in the three phases to determine what areas of weakness needed further

development and incorporation into a new supplemental instructor training program that the researcher could propose to leadership within the organization. Additionally, the researcher identified common themes in each strength or weakness explicitly related to the research questions of this study, and the researcher examined what prior research theories and literature could substantiate the reasons for the results and lay a foundation for recommending a new supplemental training program.

Trustworthiness and Authenticity

The researcher continuously used data triangulation throughout the study due to the multiple forms of data the researcher collected throughout the three research phases. However, triangulation also establishes validity and reliability as the researcher reports consistent observations in their findings, not just one observation that may be an anomaly.²¹ Additionally, triangulation proves beneficial by not just providing confirmation of findings but also more comprehensive data, increased validity, and enhanced understanding of a phenomenon.²² Therefore, the researcher constantly compared participants' performance observations and interview responses with past performance evaluations to cross-check information accuracy and credibility. This member-checking method increases validity due to the researcher ensuring they did not misinterpret the data that the participants provided by bringing the data analysis back to the participants for verification.²³ The researcher also conducted this comparison to ensure that there was an actual historical record of these deficiencies noted previously and that the annotated weakness was not just a one-time performance mistake during the observation or a complaint voiced in the interview. Specifically, the researcher cross-referenced the observations with each participant's past Form 206 from within the past year.

²¹ Creswell and Poth, *Qualitative Inquiry and Research Design*; and Ashatu Hussein, "The Use of Triangulation in Social Sciences Research," *Journal of Comparative Social Work* 4, no. 1 (April 2009): 106–17, <https://doi.org/10.31265/jcsw.v4i1.48>.

²² Abir Bekhet and Jaclene Zauszniewski, "Methodological Triangulation: An Approach to Understanding Data," *Nurse Researcher* 20, no. 2 (November 2012): 40–43, <https://doi.org/10.7748/nr2012.11.20.2.40.c9442>.

²³ Otero and Harlow, "Getting Started in Qualitative Physics Education Research."

Last, throughout the entire study, the researcher communicated and coordinated in casual conversations with the other department leads within the 128th Air Control Squadron as the SMEs in this field. The use of SMEs strengthened the dependability of the information throughout this study due to SMEs being knowledgeable in the subject matter and possessing the technical expertise that enables them to relate better to the technical needs of the learner.²⁴ The other three department leads and the researcher, as the fourth and final lead, are all considered lead instructors due to their full-time employment within the organization and generally possess a higher skill level in terms of tactical aircraft control ability. Therefore, the researcher used the department leads for peer-review to confirm the validity of collected and analyzed data to determine if the other department leads and primary instructors had noticed a historical trend in the same deficiency areas amongst the other ABM instructors.

Ethical Considerations

The researcher submitted this research to the Office of Research Compliance and Institutional Review Board (IRB) through Baylor University for review and received a nonhuman subjects research determination. Despite the researcher conducting interviews with participants, this determination was decided on due to the intrinsic research primarily focusing on the current training program and not the participants themselves. Furthermore, the researcher obtained permission from organizational leadership to conduct this study, and all participants signed written informed consent forms agreeing to partake in this study. Throughout this study, the researcher also reported and ensured that proper research and accurate data findings were in line with identified trends and spot evaluations from the 128th Air Control Squadron's standards and evaluations section because conducting ethical research and process improvement within an organization is of the utmost importance. The squadron's standards and evaluations section is responsible for conducting spot inspections or evaluations, along with annual reoccurring and required evaluations,

²⁴ Anthony P. Carnevale, Leila J. Gainer, and Janice Villet, *Training in America: The Organization and Strategic Role of Training* (San Francisco, CA: Jossey-Bass, 1990).

to ensure that all tactical aircraft controllers are qualified, certified, and ultimately keeping the aircraft under their control safe. Ethical research promotes the validity and credibility of findings and ensures that the recipients of the research, process improvement, and other organizational changes can rely on the findings to benefit their organization.²⁵ Furthermore, conducting ethical research ensures that the current research findings are of value and preemptively builds credibility for the researcher to allow for further breakthroughs in future research.

Due to conducting this study at the researcher's place of employment, the researcher paid particular attention to remaining cognizant of bias and that the researcher's role within the study site would not influence the participants' attitudes and responses, as well as the overall reliability and validity of the study. One technique that the researcher used to minimize bias was verbalizing to all participants that there were no repercussions for their responses and that all participants and their respective responses would remain anonymous to leadership. Last, due to the inherent rank structure and chain of command, a private office setting was used to prevent participants' potential fear of reprisal. Similarly, the researcher informed all participants that participation in the study was voluntary and that all data collected would remain confidential. The researcher asked all participants during data collection if the hand-written notes the researcher collected of their feedback were correct, and they had access to receive a copy of the research the researcher conducted once the study was complete to ensure that the researcher did not misrepresent their feedback. This member-checking method increases validity due to the researcher ensuring that they did not misinterpret the data the participants provided by bringing the data analysis back to the participants for verification.²⁶

For this study's research and data collection, the researcher collected data using pen and paper during the research's observational and interview phases. The researcher secured all pen and paper responses via a personal locked safe within the researcher's office. Additionally,

²⁵ Creswell and Poth, *Qualitative Inquiry and Research Design*.

²⁶ Otero and Harlow, "Getting Started in Qualitative Physics Education Research."

the researcher kept all coding and general thematic analysis notes on a password-protected computer. The findings throughout the research process and the overall privacy of each participants' comments to interview questions were protected and kept between the participants and the researcher. This protection aided the researcher in potentially acquiring more willing participants who expressed their opinions or concerns about the current training program more openly without fear of reprisal from organizational leadership.

The researcher also maintained confidentiality by assigning participants pseudonyms, with the master list of actual names maintained only within the researcher's locked office safe. This process helped participants feel as if reprisal was nonexistent, as they knew organizational leadership was not aware of the responses to each research question. This method of assigning pseudonyms is critical in a military environment where leaders or members who outrank one another can easily reprimand subordinates for their responses or specific actions.

Limitations and Delimitations

There are several underlying assumptions concerning the selected methodology and procedures. The first assumption is that this study's specific methods and procedures would reveal current deficiencies in the training program. This assumption relied on the fact that the study participants would show areas of weakness during mission observation and would also voice their genuine concerns over areas of the current training program that need improvement. This notion of no weaknesses being present or participants not providing any process improvement feedback was a significant consideration for the researcher before starting the study. However, being currently employed at the study site, the researcher was well aware of ABMs voicing their frustrations and struggles regarding not feeling prepared to serve in the instructor role.

Next, there were two significant limitations within the context of this study. The first limitation of this research is related to the study site and, as a result, the respective size of the sample population. As explained in the site selection section of this chapter, the researcher chose this site because of a vested interest in the organization and the learners, as it is the

researcher's place of employment and the only ANG air control squadron in the upper Midwest. Access to other squadrons throughout the country was not feasible for this study, and the researcher was limited to the four ANG participants who are currently ABM instructors at the 128th Air Control Squadron. Therefore, the researcher noted that a sample of this size was not large enough to draw conclusions that are representative of the other eight squadrons and the entire ANG ABM community.

The last limitation of this research is the notion of participant reliability and accuracy in response to interview questions and performance during the observational phase of research. This performance requires special consideration since the researcher conducted the observations in a military environment and, being of mid-level management status, outranks some of the participants in this case. Therefore, one must consider if each participant was truthful and honest in their responses out of fear that the researcher or other leadership would negatively impact their careers. One technique that the researcher used to minimize bias was verbalizing to all participants that there were no repercussions for their responses and that all participants and their respective responses would remain anonymous to leadership.

Throughout this study, the researcher established delimitations as they related to the research questions, the location of the study site, and the number of participants. For example, the researcher avoided research questions related to the higher-level examination of overall Air Force policies and regulations. This avoidance was due to research time restrictions and the time and protocol requirements for receiving Air Force policy approval to change training requirements for an entire air control squadron community nationwide. Therefore, the researcher developed the research questions to gain insight into the 128th Air Control Squadron, not the broader ABM community nationwide. These insights and proposed changes to the local training program were more easily achieved and implemented by narrowing the study to just this one squadron. Additionally, in phase one of the research, the researcher avoided interviewing and collecting data from every single operational Air Force career. Utilizing every career field would have drastically extended the research timeline; therefore, the researcher kept the study to include five of the most close-

ly aligned operational career fields to ANG ABMs. The following section discusses the research findings and implications by examining the specific results from the participants' observations and feedback during the semistructured interviews.

Results

Once all three phases of data collection and analysis were complete, the researcher then used triangulation between the three phases to determine and identify the three major findings of this study. These major findings or themes related to the perceptions toward the overall effectiveness of the current IUP training program among ABMs in the ANG. The first theme that surfaced, primarily from conducting semistructured interviews in the third phase of data analysis, was participants' preference to have multimodal instruction implemented in their training instead of only relying on unimodal instruction. The second major theme the researcher identified among the participants was a desire for more direct application of learning material to their actual job responsibilities, among a desire for other andragogical principles to be implemented in their training. The last major theme the researcher identified in the current training program was a lack of aircraft control effectiveness and expertise among the participants related to both Joint and tactical-level knowledge.

Cross-Phase Theme 1: Preference for Multimodal Instruction versus Unimodal Instruction

The first major theme the researcher observed across multiple data collection and analysis phases related to a current unimodal instructional technique implemented within the ANG. Likewise, throughout the phase three semistructured interviews, the researcher noted congruence among all ANG participants who desired differentiated instructional strategies and a multimethod approach to instruction within their IUP. When conducting phase one of this study, the researcher determined that the ANG relied solely on self-guided PowerPoint slides to teach its ABM instructors. Phase one allowed the researcher to determine that this use of un-

imodal instruction was specific to the ANG and not to the active duty ABM career field.

In the phase three semistructured interviews, all four participants provided feedback on the ineffective use of PowerPoint academics in the current training program. All participants described that their respective training would have been much more beneficial if they had the opportunity to learn from multiple sources. According to the participants, these sources included instructional methods such as using supplemental videos and readings within the academics provided during the academic training components of the program. Other ideas that emerged as preferred learning methods among all participants were the recommendation for using a blend of group projects, job shadowing, and practice instruction to master the instructor training content.

Cross-Phase Theme 2:

Andragogy and Direct Application of Learning Material

The second major theme that the researcher observed across all data collection and analysis phases was a desire to have the academic material required within the training program directly apply to the ABM's role as an instructor, in line with andragogical principles. Overall, throughout the study, the researcher was able to identify a general lack of other andragogical principles throughout the IUP training program as well. For example, all participants in this study agreed and shared concerns or feedback on a desire for their training to benefit them as instructors directly. The feedback the researcher received and observed across multiple phases of this study led the researcher to conclude that roughly half of the material in the training syllabus was of an administrative nature and did not directly apply to tangible results in terms of better engagement and instruction between the instructors and students. All participants expressed that their current training ultimately served no purpose if tangible results or becoming better overall instructors is not the final outcome. This feedback left all participants feeling like the program is a check-the-box process versus an actual beneficial milestone in their career development and growth.

In addition to a desire for direct application of learned material to the participants' daily mission responsibilities as an instructor, the researcher's findings illustrate a desire for other andragogical principles. For example, the self-guided PowerPoint presentations the instructors are required to study present a *sit-and-get* approach to their learning. The participants thought that taking a more active role and having a choice in how they learn the material would be much more beneficial to their learning and overall success as instructors. Each of the participants explained that they could still learn the same material if they chose to select different methods of instruction.

Cross-Phase Theme 3:

Need for Increased Joint and Tactical-Level Academics

The third and final major theme that the researcher identified across all data collection and analysis phases relate to a need for increased implementation of both Joint and tactical aircraft-related academic content within the IUP training program for ANG ABMs. While conducting the first phase of research, analyzing the active duty and ANG syllabi, evidence showed that, unlike the active duty syllabus, the ANG had no requirements centered around instructors learning the primary content that they were responsible for teaching, which consists of tactical aircraft control fundamentals. During the phase two observations the researcher conducted, a common theme emerged between all ABM within the ANG study site: every mission controller had errors in understanding tactical aircraft knowledge. Specifically, using the observation Form 206, each participant was marked down in the area of correct fighter jet targeting and intercept communication when controlling aircraft outside Air Force aircraft platforms. These knowledge gaps also consisted of various aircraft platform capabilities and limitations centered around aircraft weapons, radar systems, and engine capabilities and limitations. The same is the case for errors observed when ABM instructors taught aircraft fundamentals from other military branches besides the Air Force. This set of errors led the researcher to conclude that there is a fundamental knowledge gap related to understanding Joint aircraft capabilities and limitations.

General Themes in Relation to the Study's Research Questions

In addition to the findings identified across all three phases, the researcher also identified several other commonalities between the three phases of the study and all the study participants. First, all participants lacked a general understanding or experience in developing various assessment strategies properly. For example, while watching the crew brief, no participants assessed the crew's understanding of the mission during the observation phase. Rather than asking questions and probing the crew members for various understandings related to the mission, the participants all briefed the crew in a one-way conversation prior to the mission, with no informal assessments to gauge understanding. These strategies most closely relate to formative and summative assessment techniques used in other educational professions. Second, throughout this research, the researcher noted a general lack of understanding of the internal equipment that ABMs use to perform the live or simulated aircraft control missions. For instance, while each participant stepped to execute the mission on the control floor, there were problems with equipment related to the radios and radars not working correctly. Each participant relied on a more seasoned member to help troubleshoot the equipment and did not know how to resolve the issue themselves. These findings relate to the study's first research question aimed at identifying deficiencies in the ANG's IUP training program due to general program requirements, current training methods, and overall program structure.

In summary, the findings the researcher identified across all three phases of the research and these other less prominent commonalities between the phases and participants lead the researcher to believe that there is room for growth and improvement within the ANG's current IUP. For example, feedback from ANG participants indicated a general dissatisfaction with current unimodal instructional techniques. This identified theme directly relates to the second research question of determining participants' perceptions toward current instructional delivery methods. Additionally, due to the triangulation the researcher used across all three research phases, the researcher identified these three primary themes related to the research questions. Specifically, these three major themes all

related to issues with unimodal instruction in the current training, a lack of andragogical principles in training, and differences in the syllabi that the two Air Force components need to remedy to create unison in the training of air battle managers for the Joint environment.

Conclusion

The purpose of this study was to examine the perceptions toward and related effectiveness of the current ABM IUP training program within the ANG. To best understand ABM instructors' unique experiences and opinions, the researcher utilized a qualitative intrinsic case study methodology to gain in-depth feedback on the current program. Specifically, the researcher used a combination of artifact analysis, observations, and semistructured interviews to analyze the program. This three-phased approach to the data collection was vital for understanding the training program and the quality of the ABM instructors.

Additionally, the researcher examined the possible overuse of a single instructional method to deliver the training, whether sound adult learning principles already exist in the training, and the benefits of implementing differentiated instructional practices to inform change in the current ANG's ABM instructor upgrade training program within the 128th Air Control Squadron. The researcher analyzed the findings from the three data collection phases by coding and placing the various strengths and deficiencies in the current program into themes. These themes then enabled the researcher to understand the more significant problems within the current training and develop recommendations for a new supplemental training program that addresses the deficiencies uncovered through this research.

The results of this study will aid in improving the quality of ANG ABM instructors, decreasing flight safety incidences, and better aligning the ANG to the active duty when deployed in a Joint environment. This study's findings, however, not only pertain to the ANG's ABM community. The research findings related to proper andragogic principles in training, unimodal versus multimodal instructional strategies, and differentiated instruction principles have broader implications for other career fields, professional military education institutions, and the broader research

related to the scholarship of teaching and learning (SOTL) within the U.S. military context. In line with Sae Schatz et al., the goal of continual refinement within the military learning environment is to “produce savvy, agile, and operationally adept individuals, teams, and organizational structures.”²⁷ Without question, the preceding research study aims to do just that and leave a mark of improvement on current training practices that will be refined for the betterment of all servicemembers.

²⁷ Sae Schatz et al., “The Changing Face of Military Learning,” *Journal of Military Learning* (April 2017): 81.



CHAPTER 8 ●

Generative AI in the PME Classroom

Lessons Learned

Kristan J. Wheaton

In November 2022, OpenAI, at the time a small artificial intelligence (AI) research company, released ChatGPT, a form of generative AI, to the public. By January 2023, this unique software, with its uncanny ability to understand and respond to natural language prompts, had more than 100 million active monthly users, making it the fastest growing consumer application ever.¹

Since then, Google, Meta, and other companies have released a large number of open and proprietary large language models (LLMs) built on the back of the same generative pretrained transformer (GPT) approach used by OpenAI. As of this writing, some of the top performing models include OpenAI's most recent version, GPT-4; Google's Gemini; the AI startup Anthropic's Claude 3, and Meta's open-source LLaMA-2. All of these models allow the user to ask, or prompt, the AI with unstructured text before receiving a response the same way in text.²

These so-called "text-to-text" models are not the only kinds of generative AI currently available. Entrepreneurs and researchers have used the same logic underpinning this technology to create a wide variety of text-

¹ "ChatGPT Sets Record for Fastest-Growing User Base: Analyst Note," Reuters, 1 February 2023.

² For more on LLMs, see Alaa Youssef et al., "The Importance of Understanding Language in Large Language Models," *American Journal of Bioethics* 23, no. 10 (2023): 6-7, <https://doi.org/10.1080/15265161.2023.2256614>.

to-image models, such as Midjourney, Dalle-3, and Stable Diffusion, and text-to-video models, such as Runway Gen 2 and OpenAI's Sora. In late 2023, the Massachusetts-based startup Sona announced a text-to-music model and the text-to-3D model company Meshy launched, and in early 2024, a company called Lore Machine released a text-to-graphic novel generator. Increasingly, however, these models are converging into what are known as multimodal large language models (MLLMs), where users can put any sort of document or picture into the AI, engage in unstructured dialogue, and then have data, code, text, images, video, and music (to name a few) as output.³

The speed and scale with which these models and their associated applications have become available have forced even the most curmudgeonly of CEOs and government leaders to take notice. Mark Murphy, the head of U.S. Enterprise Software Research at J. P. Morgan, claimed in March 2024, “The advent of generative AI is a seminal moment in tech, more so than the Internet or the iPhone.”⁴ Similarly, well-known consulting firm McKinsey stated that organizations would need to completely “rewire” their businesses to take full advantage of generative AI.⁵ Goldman Sachs, in March 2023, estimated that generative AI would eventually put 300 million jobs at risk worldwide, while simultaneously adding \$7 trillion (USD) to the world's gross domestic product (GDP) by 2033.⁶ This may well be an underestimated figure. Slack, best known for its popular collaborative work platform, reported in March 2024 that, despite continued concerns about inaccurate and biased outputs, generative AI use had increased 24 percent in the last quarter alone. Slack also reported that about 80 percent of those using generative AI reported an increase in productivity.⁷

³ Davide Caffagni et al., “The (R)Evolution of Multimodal Large Language Models: A Survey,” *arXiv* (February 2024).

⁴ “Is Generative AI a Game Changer?,” J. P. Morgan, 4 February 2024.

⁵ Eric Lamarre et al., “A Generative AI Reset: Rewiring to Turn Potential into Value in 2024,” *McKinsey Quarterly*, 4 March 2024.

⁶ Joseph Briggs and Devesh Kodnani, *Global Economics Analyst: The Potentially Large Effects of Artificial Intelligence on Economic Growth* (New York: Goldman Sachs Publishing, 2023).

⁷ “New Slack Research Shows Accelerating AI Use and Quantifies the ‘Work of Work,’” *Slack* (blog), 27 February 2024.

It is no surprise, then, that along with this extremely rapid growth in capacity and capability, the U.S. Department of Defense (DOD) has taken an increased interest in the potential inherent in generative AI. As Deputy Secretary of Defense Kathleen H. Hicks put it in the foreword to the November 2023 *Data, Analytics, and Artificial Intelligence Adoption Strategy*:

Alongside industry's advancements, DoD has for years made steady and swift improvements to its data foundation and analytics capabilities: experimenting with AI through research and development, integrating these technologies into business and warfighting functions, and laying the foundation for their use at scale. As our investment, experimentation, and innovation continues and accelerates, our task now is to drive the diffusion of these technologies across the enterprise.⁸

This top-down pressure to explore the possibilities of generative AI across the full spectrum of military activities has, of course, found its way to the Service schoolhouses. Indeed, faculty and staff at professional military education (PME) institutions were among the first to begin exploring the practical applications of generative AI to military problems.⁹

The purpose of this chapter is to focus on one such effort and to extract the lessons learned from using generative AI in the PME classroom, specifically the Futures Seminar conducted annually at the U.S. Army War College (USAWC).

What Is the Futures Seminar?

The USAWC is the U.S. Army's senior Service college. While students to the college come from all over the U.S. government as well as from more than 70 allied nations, the primary focus is on educating U.S. Army lieutenant colonels and colonels to become strategic thinkers, earning a master's degree in strategic studies along the way.

⁸ *Data, Analytics, and Artificial Intelligence Adoption Strategy: Accelerating Decision Advantage* (Washington, DC: Department of Defense, 2023), 2.

⁹ John Goulette, "Army War College Professor Incorporates AI into Teaching Strategic Leaders," Defense Visual Information Distribution Service, 17 June 2022.

The Futures Seminar is a year-long elective course within the US-AWC curriculum designed to teach students how to think rigorously about the future while also contributing to the Army's understanding of issues critical to its future success. Futures Seminar students learn how to analyze futures problems and how to manage analytic teams tasked with understanding futures problems. These skills are immediately applied to a real-world question asked by senior leader sponsors from the DOD. Products developed as a result of this effort are briefed and delivered directly to the sponsor near the conclusion of the course.¹⁰

Seminar students opt-in to the seminar early in the academic year, with the course counting toward all eight required elective credits as well as for the students' strategic research requirement. In exchange for this significant time commitment, students get to select which of the projects and sponsors they would like to work for during the course of the year.

As soon as the class begins, students are sorted, according to their preferences, into small teams, each working for a different senior leader. Each team is also assigned a faculty advisor to help guide them through the process. Students begin by working with the sponsor and the sponsor's staff to help craft the research question, which is then memorialized in a terms of reference. Sponsors have included senior leaders such as General Charles A. Flynn, commander, U.S. Army Pacific; General James E. Rainey, commander, Army Futures Command; Lieutenant General Laura A. Potter, former U.S. Army G2 and now director of the Army Staff; Rachel Jacobson, assistant secretary of the Army for Installations, Energy and Environment; and Lieutenant General Dagvin R. M. Anderson, director, Joint Staff J-7, among others.¹¹

The questions these senior leaders ask of these students are futures-focused and challenging. For example, General Rainey posed the following in academic year (AY) 2023:

What are the likely applications across the continuum of human-machine teaming which will be technological-

¹⁰ John Goullette, "Army War College Futures Program Studies Emerging Issues," Army .mil, 14 June 2022.

¹¹ Goullette, "Army War College Futures Program Studies Emerging Issues."

ly feasible, militarily relevant, and ethically acceptable through 2040?

- What are the likely legal and ethical implications which may drive requirements or generate limitations? How are those likely to change between now and 2040?
- What is the future of human-machine teaming across domains with a particular emphasis on landpower?
- What are the likely determining factors for the utilization and prioritization of autonomous systems across the scope of warfighting functions?
- How will international and social norms likely evolve regarding the use of autonomous systems in combat between now and 2040?¹²

In all, the Futures Seminar has taken on 23 such projects during the last five years.

In late 2022, immediately after OpenAI launched ChatGPT, it became obvious to the Futures Seminar faculty (the author included) that generative AI had to be incorporated into the Futures Seminar. There were four primary reasons for this. First, senior leaders had long indicated that technologies like AI were precisely the ones that would ensure that the United States maintained its battlefield advantage. Now that easily accessible examples of these kinds of technology were available, it seemed a shame to waste such an opportunity to learn about them. Second, the speed with which the AI wave was growing suggested that students, senior officers in their own right, would likely have to make decisions about incorporating AI into DOD processes immediately after graduation. Third, the USAWC needed (as did all institutions in late 2022) to have some hands-on input to inform policy going forward. Finally, this was the *Futures Seminar*. If generative AI was not an appropriate object of study here, where would it be?

¹² LtCol Owen Adams et al., *Human-Machine Teaming 2030–2040: Redefining the Continuum* (Carlisle, PA: U.S. Army War College, 2023), 1.

The policy for use of generative AI within the Futures Seminar was deliberately kept simple: use it, own it, tell us about it. Faculty allowed and encouraged students to engage with a wide variety of generative AI products.¹³ Drawing not only from popular offerings such as ChatGPT and MidJourney, the students tapped into a growing body of third-party applications that leverage application programming interfaces (APIs) from these foundational models to create new capabilities and cataloged in directories such as Futurepedia and There'sAnAIForThat.¹⁴

Students, however, were also required to “own it.” Faculty required students to acknowledge which AI(s) they had used and for what purpose in their written work. In addition, it was reinforced that when they put their name on a product, *it was theirs*—that they were responsible for the facts and logic in their work, regardless of the source. The application of this policy was made significantly easier by the unique population of the class. Virtually all the students had been battalion or battalion-level commanders and were used to having a staff and taking responsibility for staff work that they approved. Administratively, the unique structure, format, and style guidance used in Futures Seminar products also made it impossible to “have an AI do the work for you.”

Finally, students not only updated faculty on AI usage during bi-weekly team meetings but also went through a detailed after action review following the presentation of the final products to their respective sponsors. Faculty in the 2024 Futures Seminar are using much the same policy, amended, of course, to take into consideration new institutional guidance that has been issued in the interim. The lessons learned outlined in this chapter are derived primarily from this feedback and the observations of the faculty.

The results of this aggressive use policy on final products were stunning. All six teams incorporated elements of AI or AI-generated products into their final findings and briefings. While each instance of AI use was clearly identified in the context of the future-focused findings of these reports, each instance also added significant value as an illustration, a

¹³ The policy is for the USAWC classroom only as defined by the course syllabus.

¹⁴ “Futurepedia: Find the Best AI Tools & Software,” Futurepedia, accessed 6 June 2024.

demonstration, or as an example. In one memorable case, students, with no technical experience and using only publicly available AIs, manage to clone their sponsor's voice, animate the sponsor's headshot, and train an open-source chatbot on both the sponsor's biography and the team's findings. As a result, at one point, the team was able to have the bot answer the sponsor's questions about the report in the sponsor's own voice and some of the sponsor's personality.¹⁵ All this was not only directly relevant to the sponsor's question, but also a powerful demonstration of how pervasive and easy to use information manipulation tools had become.

What Were the Lessons Learned?

While there were a number of lessons learned on the technical side, these are already proving ephemeral as new technologies (or at least newer versions of older technologies) emerge. Generalizable insights, those that cut across most activities and seem to have both some longevity and utility to other faculty and administrators trying to figure out how to think about generative AI in the classroom, were harder to identify. In total, there were six:

- Generative AI is neither a savior nor Satan.
- Generative AI is like a very fast, very average staff officer.
- Generative AI is better at form than content.
- Do not start with generative AI unless you are looking for "good enough."
- Good questions are becoming more valuable than good answers.
- Artificial intelligence is not going to be enough. We are going to need artificial wisdom too.

Generative AI Is Neither Savior nor Satan

The aggressive use of AI during the entirety of the class gave virtually all students a much more nuanced view of the strengths and weaknesses of the current versions of generative AI. As discussed throughout this chap-

¹⁵ LtCol Owen Adams et al., "Human-Machine Teaming 2030–2040: Redefining the Continuum" (masters diss., U.S. Army War College, 2023), 1–161.

Have You Used Generative AI?

Among every audience with which I have discussed generative AI, there are always some who have not yet tried it. It is almost impossible to have a robust discussion with someone about the pros and cons of generative AI if they have not tried at least one version of the technology. Rectifying this is one of the first things I try to do when talking to a group.

Fortunately, there are a number of easy ways to do this without costly subscriptions or even privacy-violating registrations. My personal favorite comes by way of the website Perplexity.AI.¹ Available on any internet-connected desktop computer or mobile phone, Perplexity is a third-party application that uses OpenAI's GPT series of LLMs to drive its software.

Once on the site, simply ask it a question. I usually encourage people to ask something knowable but hard, something that the asker knows the answer to or at least knows enough to know if the answer Perplexity gives back is correct. I then usually encourage people to rate the answer on a scale from 1 to 10, where a 9 or 10 is a very good, actionable answer and a 1 or 2 is a dangerous or otherwise objectionable answer.

I first used this exercise in February 2023 and, at that time, only about 25 percent of the audience had ever tried generative AI. In addition, while there were quite a few people in the crowd who rated their answer a 9 or 10, there were also a number of 1s and 2s. Since then, the percentage of people who have used generative AI has steadily gone up along with the average score. During my last presentation in February 2024, everyone in the audience had used generative AI and there were many 9s and 10s and no 1s or 2s.

ter in detail, students finished the class with a greater appreciation for the technology but also a much more cautious approach to when and where it should be used in the defense enterprise.

More importantly, students recommended overwhelmingly that generative AI be a part of the Futures Seminar going forward. In the end-of-year evaluation, students gave text-to-text models a 4.9 (out of 5) and text-to-image models a 4.2 when asked if they should continue to be taught and used in the seminar.

Anecdotally, and from a faculty perspective, it felt as if students started the class at one end of the spectrum or the other, either firmly convinced that generative AI was a universal good or, more commonly, an unreserved evil. As time progressed, familiarity did not seem to breed either contempt or attraction. Instead, students seemed to develop understanding. These results are consistent with some studies looking at the role of familiarity. In live interactions, like the

¹“Perplexity AI,” Perplexity, accessed 6 June 2024.

ones in class, familiarity does seem to lead to increased attraction.¹⁶ This effect, in turn, has likely been mitigated somewhat by military officers' "rigorous form of collective sensemaking."¹⁷ These offsetting effects seem to have left students in a place where they were more accepting of AI in general but more cautious when it came to specific applications of AI for specific problems.

Generative AI Is Like a Very Fast, Very Average Staff Officer

As mentioned earlier, the idea of "AI as a staff officer" is particularly useful in the context of a PME institution. These students know what a staff officer is and what they do. They also know the difference between a good one and a bad one or, in this case, an average one. Indeed, comments reflecting this understanding of the role of AI became increasingly common as the year progressed.

AIs alone were rarely able to "nail it" when responding to a prompt. In nearly every case, the AI output needed tweaking or fixing, sometimes significantly. They were particularly bad at seeing and managing the relationships between the moving pieces of complex problems and particularly good at doing administrivia. The AIs excelled at helping students overcome blank-page syndrome, but it almost never presented a complete answer.

The speed with which the AI generates content also tended to fool students about the competence of the AI, at least initially. Students unused to working with AIs often misinterpreted speed for competence, a common, research-supported conception.¹⁸ Extended use always led to a more sanguine assessment of the AI's capabilities. While the output was sometimes better and sometimes worse, on the whole and by the end of the year, it was considered average.

¹⁶ Harry T. Reis et al., "Familiarity Does Indeed Promote Attraction in Live Interaction," *Personality and Social Psychology* 101, no. 3 (2011): 557–70, <https://doi.org/10.1037/a0022885>.

¹⁷ Hans Hasselbladh et al., "Why Military Organizations Are Cautious about Learning?," *Armed Forces & Society* 46, no. 3 (2020): 475–94, <https://doi.org/10.1177/0095327X19832058>.

¹⁸ P. C. M. Molenaar and M. W. van der Molen, "On the Discrimination Between Global and Local Trend Hypotheses of Life-Span Changes in Processing Speed," *Acta Psychologica* 86, nos. 2–3 (August 1994): 273–93, [https://doi.org/10.1016/0001-6918\(94\)90005-1](https://doi.org/10.1016/0001-6918(94)90005-1).

Generative AI Is Better at Form than Content

Almost a corollary to the observation above, while generative AI often produces average output, it excels in its ability to do so in a multiplicity of forms.

All communications have at least two aspects: form and content.¹⁹ *Content* refers to the message, the information, that a speaker or writer hopes to get across. It could be about politics, or philosophy, or within PME institutions, rather predictably, about war.

Form is different. The medium through which authors express ideas impacts how authors communicate those ideas. For instance, the same core concepts about politics, philosophy, or war can be conveyed very differently depending on whether they are presented through a 20-minute PowerPoint presentation or a 2,000-word paper. The form chosen defines how the author communicates the message.²⁰

For example, Anthropic's Claude AI, when asked to explain the meaning of the USAWC's motto—*Prudens Futuri*—to the incoming class of students, it came back with:²¹

The Latin phrase “prudens futuri” translates to “foreseeing the future” or “provident of the future.” It refers to the wisdom and foresight required of effective leaders. Prudent leaders do not simply react to current circumstances; they have a vision for the future and proactively position their forces to capitalize on anticipated events. As Sun Tzu advised, foreknowledge enables leaders to seize opportunities and gain advantage. Prudens futuri requires gathering intelligence, conducting analysis, forecasting threats, and planning for contingencies. Leaders who excel at prudens futuri can execute bold, unconven-

¹⁹ Gideon O. Burton, “Content/Form,” *Silva Retoricae*, Brigham Young University, accessed 19 February 2024.

²⁰ Marshall McLuhan, Quentin Fiore, and Jerome Agel, *The Medium Is the Message: An Inventory of Effects* (New York: Bantam Books, 1967).

²¹ This is officially translated as “provident for the future.” For additional information, see “What Is the US Army War College Motto and Insignia?,” USAWC Library Answers, 5 December 2022.

tional strategies to defeat enemies who rely on reactionary thinking. *Prudens futuri* epitomizes the innovative, forward-looking mindset the Army must adopt to maintain strategic dominance.²²

Asking it to explain the same thing to a 13-year-old yielded the following:

“*Prudens futuri*” is a Latin phrase that means “foreseeing the future.” It’s about being wise enough to predict what might happen next. Great leaders don’t just respond to what’s in front of them, they think ahead. Imagine you’re a general planning to attack an enemy. A wise general would send scouts to spy first and learn the enemy’s weaknesses. Then the general can make a smart battle plan to sneak around and surprise them from behind. The Latin phrase teaches that real leaders don’t just react, they use their brains to see the future and stay one step ahead. That’s what “*prudens futuri*,” or “foreseeing the future,” is all about!²³

This observation that AI is better at *form* than *content* has repercussions for all educational institutions, not just PME. As a *source*, generative AI suffers from a number of problems—bias, factual hallucinations, unreproducible and unethical outcomes—but as a writing partner, able to help craft the *form* of the message, it is unparalleled.²⁴

²² Claude AI, response to “explain the meaning of the U.S. Army War College’s motto—*Prudens Futuri*—to the incoming class of students,” 2023, <https://claude.ai>.

²³ Claude AI, response to “explain the meaning of the U.S. Army War College’s motto—*Prudens Futuri*—to a 13-year-old,” 2023, <https://claude.ai>.

²⁴ According to MIT, *factual hallucinations* refer to content generated by an AI platform that provides fabricated content that appears so realistic that it appears authentic to the user. Nicolas de Bellefonds et al., “What’s Dividing the C-Suite on Generative AI?,” Boston Consulting Group, 28 September 2023.

Do Not Start with Generative AI unless You Are Looking for “Good Enough”

LLMs are trained on existing bodies of text or images. They generate their output by doing a series of complicated statistical manipulations of their training data and predicting, in essence, what will come next in a sequence of words.²⁵ There is some ability to adjust the “temperature” of an AI model, to change the level of randomness in the model to make it seem more creative.²⁶ In general, however, the model may surprise the user, but it is, for the most part, just repackaging things that have been said before.

In short, generative AI seems to do “good enough” quite well. Students and faculty routinely noticed that generative AI was quite good at anything that did not require creativity or innovation. Boilerplate, administrative, and scut work are all excellent uses of generative AI. The flip side of this was more important, however. For exercises that did require creativity or deep insight, starting with a generative AI almost always anchored the participants, making it that much harder to come up with anything new. Again, the speed and confidence with which generative AI produces average content can scuttle any deeper or more impactful reflections.

Good Questions Are Becoming More Valuable than Good Answers

It is impossible to predict where the boom in generative AIs will lead, but one thing seems certain: these tools are rapidly driving down the cost of getting a good answer very quickly. Moreover, they are likely to continue to do so for quite some time.

This has been a trend since at least the 1800s, with the widespread establishment of public libraries in the United States and United Kingdom.²⁷ Since then, improvements in cataloging, the professionalization of the workforce, and technology, among other things, have worked to drive down the cost of getting a good answer. The quest for a less expensive but still good answer accelerated, of course, with the introduction of the

²⁵ “What Are Large Language Models (LLMs)?,” IBM.com, accessed 6 June 2024.

²⁶ Markus Karlsson, “Generative AI, Temperature and Hallucinations,” Affino, 2 June 2023.

²⁷ “A History of US Public Libraries,” Digital Public Library of America, accessed 6 June 2024.

World Wide Web in the 1990s, driving down the cost of answering even tough questions.²⁸ While misinformation, disinformation, and social media will continue to lead many people astray, savvy users are better able to find consistently good answers to harder and more obscure questions than ever before.

If the internet accelerated this historical trend of driving down the cost of getting a good answer, the rollout of generative AI to the public in late 2022 threatens to drive costs to near zero. Hallucinations and bias aside, the simple truth is that generative AI is, more often than not, able to give pretty good answers to an awful lot of questions and it is currently free or cheap to use.

All this begs the question, though: If answers are getting better, cheaper, and more widely available at an accelerating rate, what is left? In other words, if no one needs to pay for answers anymore, what can PME offer? Where is the value-added? It may well be that as the cost of a good answer goes down, the value of a good question goes up. In short, the winners in the coming AI wars are going to be the ones who can ask the best questions at the most opportune times. Indeed, becoming better at asking useful questions of AIs seems to be the whole basis of the new career field of “prompt engineering.”²⁹

The idea that asking the right question is not only smart but essential is an old one:

“A prudent question is one-half of wisdom.”

~ Francis Bacon³⁰

“The art of proposing a question must be held of higher value than solving it.”

~ Georg Cantor³¹

²⁸ “A Short History of the Web,” CERN, accessed 6 June 2024.

²⁹ “What Is Prompt Engineering?,” IBM.com, accessed 6 June 2024.

³⁰ Francis Bacon, *Index of Project Gutenberg Works of Francis Bacon*, comp. David Widger (Fairbanks, AK: Project Gutenberg, 2019).

³¹ J. J. O'Connor and E. F. Robertson, “Georg Ferdinand Ludwig Philipp Cantor,” School of Mathematics and Statistics, University of St. Andrews, Scotland, October 1998.

“If you do not know how to ask the right question, you discover nothing.”

~ W. Edwards Deming³²

Educators often think that they are already doing this; that at least one purpose of all education, certainly of higher education, is to teach students how to think critically and how, in essence, to ask better questions. But is that really true? Virtually the whole education system is structured around evaluating the quality of student answers. Educators may think they educate students to ask probing, insightful questions, but they almost always grade, promote, and celebrate students for the number of answers they get right.

What would a test, based not on the quality of the answers given but on the quality of the questions asked, even look like? What criteria would an instructor use to evaluate a question? How would a professor create a question rubric? There is very little research on this “question of questions.”³³ There are some hypotheses to be sure, but statistically significant, peer-reviewed research is thin on the ground. This represents an opportunity, of course, for intellectual overmatch for PME institutions.³⁴ If there is very little real research in this space, then any meaningful contribution is likely to move the discipline forward significantly. If what you ask in the AI-enabled future really is going to be more important than what you know, then such an investment seems not just prudent, but one with a very high potential return on even a modest investment.

Artificial Intelligence Is Not Going to Be Enough

We are going to need artificial wisdom too. AI began, as a scientific discipline, in 1956. Since then, the field has undergone two “AI winters,”

³² Demings philosophy on knowledge and information can be found in W. Edwards Deming, *The New Economics: For Industry, Government, Education* (Cambridge, MA: MIT Press, 2018).

³³ Brian Manning and Kristan J. Wheaton., “Making ‘Easy Questions’ Easy: The Difficulty of Intelligence Requirements,” *International Journal of Intelligence and CounterIntelligence* 26, no. 3 (2013): 597–611, <https://doi.org/10.1080/08850607.2013.780562>.

³⁴ *Developing Today's Joint Officers for Tomorrow's Ways of War: The Joint Chiefs of Staff Vision and Guidance for Professional Military Education & Talent Management* (Washington, DC: Joint Chiefs of Staff, 2020).

during which funding almost completely dried up, and a number of other lulls, during which both researchers and funding agencies questioned whether anything would ever come of the effort.³⁵

Yet, nearly 70 years later, it is obvious that intelligence is not enough. While generative AI and its associated technologies are undoubtedly transformational, they are unlikely, by themselves, to create a wise leader, one of the goals of most PME institutions. Recent research into both the definitional and neurological bases for wisdom suggest the reasons why. Specifically, wisdom appears to be an emergent property derived from nine distinct sub-processes:

- general knowledge of life and social decision making
- emotional regulation
- prosocial behaviors like compassion and empathy

³⁵ Rockwell Anyoha, “The History of Artificial Intelligence,” *Science in the News* (blog), Harvard.edu, 28 August 2017.

How Good Are Your Questions?

Imagine you have told a group of students that they are going to pretend they are about to go to a job interview. They know, as with most interviews, that once the interview is over, they will be asked, “Do you have any questions for us?” You task the students to come up with interesting questions to ask the interviewer.

Here is what you get from the students:

- What are the biggest challenges that I might face in this position?
- What are the next steps in the hiring process?
- What is different about working here than anywhere else you have ever worked?

What do you think? Which question is the most interesting? Which question gets the highest grade? If you are like the vast majority of the people I have asked, you say the third question. But why? Sure, you can come up with reasons after the fact, but where is the research that indicates why an interesting question is interesting? We are left, like Supreme Court Justice Potter Stewart and the definition of pornography, with “I know it when I see it.”¹

¹ *Jacobellis v. Ohio*, 378 U.S. 184 (1964).

- insight or self-reflection
- acceptance of different value systems
- decisiveness
- spirituality
- openness to new experiences
- sense of humor³⁶

Generative AI can pretend to be decisive or funny or empathetic, but, because it does not understand the consequences of its actions, and is not programmed to care, that is all generative AI can do—pretend. Fortunately, there are researchers who are beginning to explore the concept of artificial wisdom and its associated field of machine common sense.³⁷

Where Will Generative AI in PME Institutions Go Next?

While the rapid improvement in generative AI models will eventually slow and a consolidation of both providers and application developers is all but inevitable, there is still a good bit of room for growth in everything touching on this novel and exciting technology. PME institutions will be unlikely to have much say in how this technology develops. Instead, faculty and students will be tracking each new version, each new software package, in real time.

There are three strategies to pursue when technological disruption makes change the only constant:³⁸

- Beat them: invest aggressively in the disruptive technology. In the case of PME institutions, this would mean carving out time in the curriculum to teach fundamental concepts such as prompt engineering. At a minimum, it

³⁶ Dilip V. Jeste and Ellen E. Lee, “The Emerging Empirical Science of Wisdom: Definition, Measurement, Neurobiology, Longevity, and Interventions,” *Harvard Review of Psychiatry* 27, no. 3 (May/June 2019): 127–40, <https://doi.org/10.1097/HRP.0000000000000205>.

³⁷ Dilip V. Jeste et al., “Beyond Artificial Intelligence: Exploring Artificial Wisdom,” *International Psychogeriatrics* 32, no. 8 (2020): 993–1001, <https://doi.org/10.1017/S1041610220000927>; and Howard Shrobe, “Machine Common Sense (MCS),” DARPA.mil, accessed 6 June 2024.

³⁸ Joshua S. Gans, “Keep Calm and Manage Disruption,” *MIT Sloan Management Review*, 2 February 2016.

strongly suggests that aggressive monitoring of the AI space is warranted.

- **Join them:** acquire the necessary expertise in the disruptive technology. The simplest way to do this for most PME institutions is to include AI expertise in new faculty and staff hiring announcements. Coupled with ongoing faculty development sessions, this should allow such institutions to acquire the requisite knowledge.
- **Wait them out:** identify and leverage key complementary assets. PME institutions likely have a much better grasp of things like wisdom and the right questions to ask than AI developers. This knowledge has value but typically requires some resources to fully develop and package. Accomplishing such a mission would not only be worthwhile but would also help ensure that military equities are well-represented in what is highly likely to be a commercial-focused industry.

As generative AI develops and the landscape becomes more stable, other, more refined strategies and lessons learned might well come into focus. In the meantime, anecdotal evidence and simple surveys might still provide the necessary cues to take future steps with confidence.

Summary and Conclusion

This chapter explored the lessons learned from integrating generative AI into the Futures Seminar at the U.S. Army War College. By encouraging students to engage with AI tools as part of their futures research projects, several key insights emerged: AI is neither a panacea nor a peril, but rather a tool akin to a fast but average staff officer; AI excels more at form than content; AI is best suited for “good enough” tasks rather than those requiring deep creativity; the ability to ask good questions is becoming paramount; and artificial wisdom, not just narrow intelligence, is likely the next step to truly human-like performance across the broadest possible range of activities.

There are some limitations to these findings, however. The Futures

Seminar represents a unique educational context within PME, focused on strategic foresight and sponsored research projects. The lessons learned may not fully translate to other PME settings or to undergraduate or graduate education more broadly. Additionally, the rapid pace of change in generative AI capabilities means that some of the specific observations about the strengths and weaknesses of the technology will quickly become dated.

Nonetheless, this chapter aims to contribute to the Scholarship of Teaching and Learning (SOTL) by offering an early case study of integrating a disruptive new technology into the classroom. As Kathy Quardokus Fisher and G. Alex Ambrose of Notre Dame Learning define it, SOTL involves “a systematic approach to collecting and analyzing data related to your teaching practice and sharing that work with others.”³⁹ By rigorously examining how generative AI impacted student learning in the Futures Seminar, this chapter embodies that spirit of inquiry.

More specifically, within military educational contexts, this chapter aligns with calls to enhance the intellectual agility of military leaders in the face of accelerating technological change.⁴⁰ It also exemplifies the “active learning” approaches advocated by the *Journal of Military Learning* as essential for developing the critical and creative thinking skills needed in contemporary military leaders.⁴¹

In sum, while generative AI presents both opportunities and challenges for PME, engaging in structured experimentation and reflection on its implementation, as this chapter has done, will be essential for harnessing its potential to enhance student learning and develop future-ready leaders. Continued SOTL research will be invaluable for building a knowledge base around the most effective practices for leveraging AI in PME.

³⁹ “Scholarship of Teaching and Learning: What, Why, & How,” Notre Dame Learning, accessed 6 June 2024.

⁴⁰ *Army Leadership and the Profession*, Army Doctrine Publication No. 6-22 (Washington, DC: Department of the Army, 2019).

⁴¹ Mark Hamilton, “Prioritizing Active Learning in the Classroom,” *Journal of Military Learning* (October 2019): 3–9.

Part 3

ASSESSMENT



CHAPTER 9

Enhancing Online PME Assessment

Design-Based Research for Rubric Development
and Evaluation in the Global College
of Professional Military Education

Robert S. Nyland, PhD; and Carl P. Watts, PhD

Introduction

Rubrics play a crucial role in higher education, as they ensure the use of transparent, consistent, and fair performance standards. Rubrics are particularly important for the Global College of Professional Military Education (GCPME) at Air University. With more than 250 full-time and adjunct faculty involved in grading thousands of courses every year, the use of rubrics is fundamental for promoting reliable and consistent assessment. This chapter outlines the structure and results from a quality improvement and evaluation study conducted at GCPME to test newly developed essay and discussion rubrics. This study serves as an exemplary case of iterative design-based research for enhancing educational products.

Megan J. Hennessey makes a convincing case for applied research in the professional military education (PME) environment. She observes that PME institutions can learn from civilian counterparts (a point endorsed in this chapter), but there are limits to this because PME is a unique environment. Hennessey therefore suggests that “professional military education scholars, administrators, and educational methodologists should incorporate frequent, methodologically rigorous interven-

tions—or experiments—in the classroom to test educational strategies.”¹ The GCPME rubric trials answered this call. The authors conducted a rigorous sequence of rubric trials using multiple GCPME courses in real time but proceeded cautiously to meet a key standard: “Students’ attainment of learning outcomes should not be put at risk by educational experimentation.”² Indeed, the rubric trials were designed to *enhance* student attainment of course learning outcomes by using rubrics as an instructional scaffold, and not just a mechanism of assessment.³ Hennessey observes that “to lead to real change, resulting analyses and findings of applied educational research should be shared publicly.”⁴ The authors are pleased to share the results of the GCPME rubric trials here because it will reach the intended audience in the PME community.

This chapter first describes the institutional context of GCPME. Second, it briefly explains why rubric development has garnered more attention from academics in U.S. higher education and PME institutions during the last few years. This is essential because, as Stase L. Wells has commented, rubrics “serve as evidence that graduate-level JPME institutions recognize the importance of feedback in teaching and learning processes.”⁵ Third, the chapter discusses the origins and purpose of the GCPME rubric trials, which predated the emphasis on rubrics in the Joint Chiefs of Staff guidance for outcomes-based military education (OBME).⁶ Fourth, the chapter outlines the research methods adopted during the rubric trials and provides a detailed account of the iterative approach to

¹ Megan J. Hennessey, “How Can We Know if Professional Military Education Works?,” *War on the Rocks*, 3 August 2020.

² Hennessey, “How Can We Know if Professional Military Education Works?”

³ Kiruthika Ragupathi and Adrian Lee, “Beyond Fairness and Consistency in Grading: The Role of Rubrics in Higher Education,” in *Diversity and Inclusion in Global Higher Education: Lessons from Across Asia*, ed. Catherine Shea Sanger and Nancy W. Gleason (New York: Palgrave Macmillan, 2020), 73–95.

⁴ Hennessey, “How Can We Know if Professional Military Education Works?”

⁵ Stase L. Wells, “Building Trust and Success Through Dialogic Feedback in Joint Professional Military Education,” in *Developing Military Learners’ Communication Skills Using the Scholarship of Teaching and Learning*, ed. Megan J. Hennessey (Maxwell AFB, AL: Air University Press, 2022), 58.

⁶ “Guidelines for Outcomes-Based Military Education Assessments,” in *Outcomes-Based Military Education Procedures for Officer Professional Military Education* (Washington, DC: Joint Chiefs of Staff, 2022), encl. D.

rubric refinement. The fifth section of the chapter presents some key lessons. The most important finding is that rubrics are a necessary but insufficient mechanism for promoting reliable and consistent assessment. PME institutions are more likely to achieve this objective—now laid down as a required OBME standard—if they adopt similar quality assurance systems to those used in civilian public examination systems in the United States and the United Kingdom.⁷

About the Global College of Professional Military Education

The GCPME is one of the colleges at Air University, which is institutionally accredited by the Southern Association of Colleges and Schools Commission on Colleges (SACSCOC) and is physically located on Maxwell Air Force Base in Montgomery, Alabama. The GCPME mission is to “develop and teach relevant, quality curriculum and deliver a high-value student experience.” The GCPME offers seven programs asynchronously via distance learning to 36,000 students distributed around the world. Three of these programs serve the enlisted PME (EPME) community of approximately 9,400 students. Airman Leadership School is a 60-hour program that develops senior airmen by strengthening their knowledge about leadership and followership while broadening their understanding of the missions and cultures of the U.S. Air Force. The Noncommissioned Officer (NCO) Academy is a 65-hour program that enhances the development of NCOs by reinforcing their knowledge about leadership and management while increasing their level of understanding of joint missions and cultures. The Senior Noncommissioned Officer Academy is also a 65-hour program that enriches the growth of senior NCOs by providing concepts to enable them to operate at an Air Force organizational level or alongside other elements of the U.S. military.⁸

A further four programs serve the officer PME community of about 26,500 students. Squadron Officer School is a 180-hour primary developmental education (PDE) program that prepares captains for greater responsibilities at the squadron level while also building critical thinking

⁷ “Guidelines for Outcomes-Based Military Education Assessments.”

⁸ “Global College of PME,” Air University, accessed 25 September 2024.

skills in the application of airpower for national security goals and objectives. Air Command and Staff College (ACSC) is a 180-hour intermediate developmental education (IDE) program and Joint professional military education (JPME) Phase I program that takes 260 hours to complete. It prepares field-grade officers to serve in operational-level command or staff positions by emphasizing the application of airpower in Joint campaign planning at the operational level of war. Air War College is a senior developmental education program that requires 200 hours to complete. It provides a strategic, “air-minded” curriculum that prepares graduates to provide strategic leadership, appropriate expertise, and critical thinking in support of national security objectives. Finally, the Online Master’s Program (OLMP) is a 30-credit degree that combines the subject matter of ACSC with rigorous research-based inquiry. The program fosters the development of higher-order thinking by challenging students to think critically about applying air and space power in Joint campaigns at the operational level of war. The curriculum addresses topics including contemporary Air Force operations, national security, leadership, and Joint warfare. The OLMP is offered in four concentrations: Joint warfare, leadership, nuclear weapons, and operational warfare. Students who complete the Joint warfare concentration also earn JPME Phase I credit and IDE credit on their Air Force record.⁹

Rubrics and PME

Until about a decade ago, “mentioning rubrics to faculty members in many fields may have brought forth looks of confusion, consternation, or disinterest.”¹⁰ Interest in rubrics and assessment was confined largely to colleges of education, and indeed this was the context in which one of the authors of this chapter became familiar with this important but relatively neglected aspect of teaching and learning. More recently, accrediting bodies have encouraged or required higher education institutions to address their use of rubrics, and many have drawn on rubrics for undergraduate

⁹ “Global College of PME.”

¹⁰ Kevin Schoepp, Maurice Danaher, and Ashley Ater Kranov, “An Effective Rubric Norming Process,” *Practical Assessment, Research, and Evaluation* 23, no. 1 (2018): 1, <https://doi.org/10.7275/z3gm-fp34>.

education developed by the American Association of Colleges and Universities.¹¹ Within the PME community, there is a paucity of literature on rubrics.¹² This has left PME institutions poorly equipped to deal with the sudden focus on rubrics thrust forward by the exigencies of OBME. The J7 Directorate for Joint Force Development is holding PME institutions accountable for developing rubrics that “provide a consistent basis for assessing performance . . . and allow for the collection of assessment results for evaluation and applied for continuous improvement.”¹³ The GCPME rubric trials originated *before* the OBME imperative became clear. As this chapter unfolds, it will show how OBME requirements changed the GCPME direction of travel in its rubric development. The reasons that impelled the rubric trials and their initial focus are discussed below.

The GCPME Rubric Trials: Origins and Purpose

The GCPME benefits from a strong instructional design and development team that maintains a focus on quality assurance using a course design quality checklist. Assessment is the fifth item in this checklist and specifies that the institution “reviews the methods, frequency, and curricular alignment of the course’s assessments, as well as the quality of its grading rubrics [and] feedback opportunities.”¹⁴ Until 2021, the chief focus was on the alignment between assessment, course learning outcomes, and lesson objectives. At that juncture, the authors of this chapter stood up a rubrics working group (RWG) to examine potential solutions to some of the longstanding issues with GCPME rubrics to which instructors, course directors, and administrators had often drawn attention. These included the number of assessment criteria, high point values attached to

¹¹ “Valid Assessment of Learning in Undergraduate Education (VALUE),” American Association of Colleges and Universities, accessed 25 September 2024.

¹² Rubrics are barely mentioned in PME books about communication and assessment. For example, the term *rubric* appears three times in Hennessey, *Developing Military Learners’ Communication Skills*, and just twice in Einike Pilli and Tiia-Triin Truusa, *Assessment in Professional Military Education* (Tartu, Estonia: University of Tartu Press, 2021).

¹³ “Guidelines for Outcomes-Based Military Education Assessments.”

¹⁴ Stephanie Teague Hostetter, “Developing a Quality Assurance Approach for an Online Professional Military Education Institution,” *Journal of Applied Instructional Design* 11, no. 2 (2022), <https://doi.org/10.51869/112/sth>.

criteria, inappropriate point ranges in the performance levels, lack of clarity in performance descriptors, excessive weighting attached to communication criteria, and distortions to student grades created by awarding points for meeting basic criteria such as timely submission and correct formatting. In general, course directors and faculty did not consider the rubrics to be user friendly or accurately reflect student performance.

Given this litany of concerns, it is legitimate to ask why problems persisted for so long. In part, this reflected the generalized PME failure to prioritize rubric development discussed above. However, in the case of GCPME, there were also some institutional and structural factors that militated against a sustained focus on rubrics, including the lack of a faculty member with experience and interest in rubric development, rapid turnover of personnel, the high operational tempo of an online college that operates all year round, a need to meet multiple higher headquarters directives regarding various aspects of the PME curriculum (e.g., learning outcomes with respect to China), the introduction of the online EPME program, and multiple lines of effort to improve teaching and learning (e.g., introduction of a social learning model driven by integration of several new technologies). In that context, the RWG had to muster and maintain significant energy and focus to overcome the inertia that had left flawed rubrics in place for more than a decade.

The RWG drafted multiple rubrics for testing and iterative trials took place in the OLMP and IDE programs during 2022 and 2023. Most of the selected courses fell within the remit of the Strategy and Securities Studies Department, in which one of the authors worked. Additional courses from other departments were included to socialize the rubric's initiative more broadly within the organization. The trials utilized a live comparison quality improvement study between equivalent courses employing the rubrics and those that did not. During these tests, several questions were used to guide the evaluation of the rubrics. First, how do student grades differ between the existing rubric and the new rubric? Second, how does the new rubric impact the student writing process? Third, how does the new rubric influence interrater reliability? GCPME leadership advised the RWG to proceed cautiously and incrementally so that stu-

dent achievement would not be adversely affected by the experimental approach. The design-based research methodology of the trials—which has been underused in the PME community—is discussed in the next section of this chapter.

Methods

Design-Based Research

This quality improvement and evaluation project utilized the framework of design-based research (DBR), which is an “approach that engages in iterative designs to develop knowledge that improves educational practices.”¹⁵ DBR was derived from scholars, including Ann L. Brown and Allan Collins, who advocated for the concepts of “design experiments” and “design research.”¹⁶ Brown described herself as a “design scientist,” someone who simultaneously embodied the roles of designer of the learning environment and educational researcher.¹⁷ In contrast to the laboratory environment in which many natural science research studies are conducted, design-based research “occurs in the buzzing, blooming confusion of real-life settings where most learning actually occurs.”¹⁸ The messy nature of learning environments requires those engaged in DBR to collect as much contextual data as possible so that they can identify elements that contribute to the success or failure of a design.¹⁹

From a process perspective, DBR engages in “progressive refinement,” wherein designs are iteratively improved.²⁰ To make the process more

¹⁵ Matthew Armstrong, Cade Dopp, and Jesse Welsh, “Design-Based Research,” in *The Students' Guide to Learning Design and Research*, ed. Royce Kimmons and Secil Kaskurlu (Provo, UT: EdTech Books, 2020), 37.

¹⁶ Ann L. Brown, “Design Experiments: Theoretical and Methodological Challenges in Creating Complex Interventions in Classroom Settings,” *Journal of the Learning Sciences* 2, no. 2 (1992): 141–78, https://doi.org/10.1207/s15327809jls0202_2.

¹⁷ Brown, “Design Experiments.”

¹⁸ Sasha Barab and Kurt Squire, “Design-Based Research: Putting a Stake in the Ground,” *Journal of the Learning Sciences* 13, no. 1 (2004): 4, https://doi.org/10.1207/s15327809jls1301_1.

¹⁹ Allan Collins, Diana Joseph, and Katerine Bielaczyc, “Design Research: Theoretical and Methodological Issues,” *Journal of the Learning Sciences* 13, no. 1 (2004): 15–42, https://doi.org/10.1207/s15327809jls1301_2.

²⁰ Collins, Joseph, and Bielaczyc, “Design Research,” 18.

concrete and operational, Susan McKenney and Thomas Reeves drew from existing research to compose a generic model for DBR, consisting of three main phases:²¹

- Analysis and exploration: team members work to understand the problem in question and potential existing research and theories that would inform the development of a solution.
- Design and construction: the learning intervention or environment is designed to address the problem identified in the analysis and exploration phase. Additionally, care is given to appropriately instrumenting the learning and research environment so that data can be used to inform the evaluation and reflection phase.
- Evaluation and reflection: data from implementation is collected and analyzed. Lessons are taken from this phase to inform future iterations of designs and research.

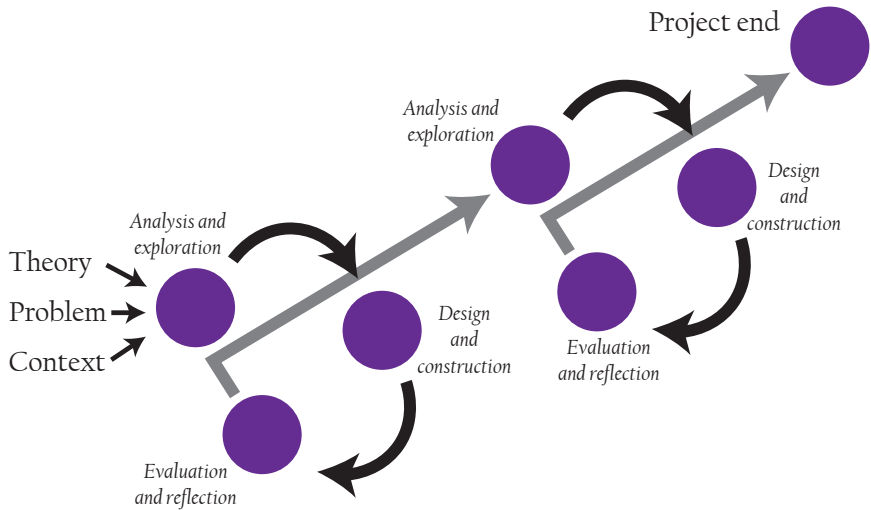
A representation of this three-phase model can be seen in figure 1. This shows that researchers cycle through the phases as they progressively refine the design to solve the problem and inform the theory being investigated.

While DBR has been used in various contexts, its impact on military education and training in the United States and other countries is nascent. A U.S. Department of Education study recommends that DBR should be used in military and training environments, especially given the existing use of virtual games and simulations.²² Antti Rissanen and Kalle Saastamoinen have provided a case study of how they used DBR to design curriculum for six courses at the National Defence University

²¹ Susan McKenney and Thomas Reeves, *Conducting Educational Design Research* (New York: Routledge, 2012), 76–82.

²² Marie Ann Bienkowski, Mingyu Feng, and Barbara Means, *Enhancing Teaching and Learning Through Educational Data Mining and Learning Analytics: An Issue Brief* (Washington, DC: Office of Educational Technology, Department of Education, 2012).

Figure 1. The iterative process of design-based research



Source: Recreated by MCUP from Armstrong, Dopp, and Welsh, “Design-Based Research,” 39.

of Finland.²³ Another case study looked at the use of DBR to create the “Total Learning Architecture,” an educational technology specification developed by the U.S. Department of Defense (DOD) Advanced Distributed Learning (ADL) organization. In a recent doctoral dissertation, Karl Miehl reports on the outcomes of using DBR to create a professional development program for civilians working in the Naval Air Warfare Center Training Systems Division.²⁴ The current report on the iterative design of the GCPME rubric trials contributes to the limited body of DBR done in the PME community.

Project Mechanics

Conducting the rubric trial involved coordination among GCPME leader-

²³ Antti Rissanen and Kalle Saastamoinen, “A Design-Based Approach to Course Planning and Development,” *Education Reform Journal* 3, no. 2 (2018): 78–87, <http://dx.doi.org/10.22596/erj2018.0302.78.87>.

²⁴ Karl Miehl, “Design-Based Research toward a Professional Development Program” (PhD thesis, University of Central Florida, 2021).

ship, course directors, course instructors, learning architects, instructional designers, and learning management system administrators. Because of the operational tempo and scale of courses in the GCPME, the team was careful and intentional about implementation timing and coordination with the various stakeholders. While every part of the process is not detailed here, the following are critical milestones that occurred in each iteration of the trial:

- Course selection: at least a month before each trial, the project team determined which of the courses were included in the trial. This selection process involved various factors including course cadence and instructor support and bandwidth. Once course selection had been determined, the team initiated necessary processes to build the new rubrics into the courses.
- Instructor selection: several weeks before each trial began, the team worked with course directors to identify which instructors would be good candidates to trial the new rubrics. Course directors recommended potential instructors based on their experience teaching the course, adaptability to course changes, and perceived likelihood of providing robust and candid feedback about the rubrics.
- Instructor onboarding: once instructors were selected, they received a memorandum about the rubric, including its purpose and expectations for feedback. Additionally, the project team requested that instructors send out a precomposed announcement about the trial.
- Student onboarding: in the announcement, students were introduced to the rubric used in their course. This included an overview of its purpose and a request for feedback at the conclusion of the course (via survey link).
- Data collection: at the end of the trial term, the team collected data through the methods described in the next section.

Data Collection

To support this quality improvement and evaluation project, the project team aimed to gather multiple forms of data that would help assess impacts on performance and gather the perspectives of users. Though the specifics of the data slightly differed across iterations, the following four primary sources were used.

Student Perceptions

In the classes that used the new rubric, student perceptions were collected via a survey link sent out by the instructor in an end-of-term announcement. These voluntary surveys were anonymous, short, and collected no demographic information. Survey questions were modified slightly during each iteration to collect information specific to that version of the rubric.

Instructor Perceptions

Like students, instructor perceptions were collected via an anonymous survey. The survey link was sent to the instructor after the term had ended so that they had completed their experience of using the new rubric for grading.

Grades

At the end of the term, grades were collected by a research group member. Grades were collected only for those assignments that used the rubric. For comparison, grades were also collected for nontrial rubric sections of the same courses that ran at the same time. Identifying information was removed and data was analyzed using the Python programming language and associated open-source packages.²⁵ The pandas package was used to clean and load data into a tabular format, which allowed for easy summarization and visualization of data, while the pingouin package was used to analyze the data for significant differences between group grades.²⁶

²⁵ “Python Language Reference, version 3.10,” Python Software Foundation, accessed 26 April 2024.

²⁶ Wes McKinney, “Pandas: Powerful Python Data Analysis Toolkit,” version 1.4.2, Python Software Foundation, accessed 26 April 2024; and Raphael Vallat, “Pingouin,” version 0.5.1, Python Software Foundation, accessed 26 April 2024.

Interrater Reliability Exercises

To assess unmediated interrater reliability (without rubric norming), a sample of graded final essays were retrieved from each of the selected trial rubric sections. The anonymized graded essays were then provided to another instructor teaching the same course participating in the rubric trial. After the second instructor completed their grading, the assignment scores were compared.

Participants

The rubric trials ran in several courses in both the ACSC (IDE) program and the OLMP. Facilitated courses in the ACSC program are between two and three weeks in duration and run every month. Courses in the OLMP are eight weeks in duration and are taught in specific terms that run through the year. Table 1 shows the timing, courses included, and number of course offerings.

Each of the following results sections outlines the intended purpose of each iteration of the rubric, the results that were gathered via the data collection process for that iteration, and lessons learned that were brought forward to the next iteration of the rubric.

Summer B 2022, 50-point Essay Rubric

Purpose and Structure

The original GCPME essay rubric used 325 points distributed across five assessment criteria: analysis (100 points), course concepts and support (100 points), writing style (50 points), writing mechanics (50 points), and administrative (25 points). For the Summer B OLMP trial, the RWG modified the essay rubric in several ways. First, it eliminated the administrative criterion. Second, it created a unified communication criterion to replace writing style and writing mechanics. Third, it added two new structural criteria for the introduction and conclusion. Fourth, it reduced the total number of rubric points from 325 to 50. Fifth, the achievement level descriptors (e.g., excellent, outstanding) were replaced with performance-level indicators (1, 2, 3, 4, 5).

The trial essay rubric used 50 points also distributed across five assessment criteria: introduction (5 points), analysis (15 points), support

Table 1. Rubric trial timing and data collection methods

Term	Program	Courses (number of sections)	Data collection
Summer B 2022	OLMP	JA-5510 (1) LC-5510 (2) NS-5510 (2)	Student survey Instructor survey Grades Interrater reliability
Spring A 2023	OLMP	AP-5510 (2) NS-5510 (3) WS-5510 (4)	Student survey Instructor survey Grades
March 2023	ACSC	ISS-602A (5) LDR-602A (4) AIR-603A (4)	Student survey Instructor survey Grades
June 2023	ACSC	ISS-602A (4) LDR-602A (4) AIR-603A (4)	Student survey Instructor survey Grades

Source: courtesy of the authors.

(15 points), conclusion (5 points) and communication (10 points). The RWG had several goals for this initial version of the trial essay rubric, some of which were focused on improving student performance, and others that were designed to condition instructor use of the rubric. The elimination of administrative points was intended to address the fact that students received up to seven percent of their essay score for meeting simple expectations, including timely submission, and correct formatting, which the RWG recognized to be inappropriate in graduate-level education. The unified communication criterion constituted an appropriate weighting of 20 percent, in contrast with the previous weighting of 31 percent attached to writing style and writing mechanics. The new criteria for the introduction and conclusion were intended to promote a focus on structural features of student writing that are often analytically weaker (and occasionally missing). The number of points in the rubric was reduced drastically because the RWG believed that a narrower point range in each criterion would result in better interrater reliability. The traditional performance labels were jettisoned because the RWG thought that

this would lessen instructor tendencies toward “intuitive” grading and encourage students to reflect more on their scores, by promoting a greater focus on the descriptors in the rubric.

Grades

Grades for each of the essay assignments were extracted from both the traditional and rubric trial sections of the course. The specific number and placement of essays depended on the scheme of assessment in each course. Average scores for each essay are displayed in table 2. Independent sample t-tests were run to compare scores between the sections of the courses using the trial rubric and those using the traditional rubric. In most instances, scores for the essays assessed using the trial rubric were significantly lower than the scores for essays assessed using the traditional rubric. However, the difference reduced over time, pointing to a potential acclimation effect for using the new rubrics.

Interrater Reliability

To investigate interrater reliability using the trial rubric, the team conducted an exercise after instructors finished grading. In this exercise, instructors were provided with a sample of student papers from the other corresponding section of the same course that also used the trial rubric. Instructors were then asked to grade the sample using the rubric. The scores provided from the other instructor were then compared with the essay score that was given to the same student paper by the original instructor. Note that this exercise was not completed for JA-5510 because only one of the course sections used the trial rubric. Table 3 shows the average grade differences between student papers overall and for each criterion in the rubric.

The table shows that there was often a large difference in scores between the instructors, with as much as a 10-percent variation between the scorers. While this project did not have a baseline to compare these differences against, it suggests that interrater reliability was still an issue with the trial rubric.

Table 2. Grade differences for Summer B 2022 trial

Course	Essay 1			Essay 2			Essay 3		
	Trial	Traditional	Diff	Trial	Traditional	Diff	Trial	Traditional	Diff
JA-5510	70.0% (n= 16)	89.5% (n=48)	-19.5**	89.5% (n= 16)	89.5% (n=48)	0			
LC-5510	83.1% (n=28)	91.8% (n=28)	-8.7**	90.1% (n=28)	92.2% (n=28)	-2.1			
NS-5510	87.0% (n= 28)	90.9% (n=40)	-3.9*	85.3% (n= 28)	91.4% (n=40)	-6.1**	84.8% (n= 28)	89.8% (n=40)	-5*

** p < .01 * p < .05

Source: courtesy of the authors.

Table 3. Mean grade differences in interrater reliability exercise.

Course	Instructor	Overall Δ	Introduction Δ	Analysis Δ	Support Δ	Conclu- sion Δ	Communi- cation Δ
LC-5510	1	-1.6	-0.4	-1.0	0.4	-0.6	0.0
	2	5.6	0.2	2.2	1.6	0.2	1.4
NS-5510	1	3.5	0.7	0.5	0.7	0.1	1.5
	2	6.1	-0.3	2.4	3.1	0.5	0.4

Source: courtesy of the authors.

Instructor Perceptions

Once instructors completed grading, they were asked to provide feedback about using the rubric via an anonymous survey to which four of the five instructors responded. Instructors were asked to respond with agreement to various statements in the survey. Overall, they found the rubric easy to use (75 percent strongly agree, 25 percent agree), that it helped to focus on the most important parts of student writing (75 percent strongly agree, 25 percent agree), that it was a good way to assess student work (75 percent strongly agree, 25 percent agree), and that it was helpful in guiding feedback on student writing (50 percent strongly agree, 25 percent agree, and 25 percent neutral). All the instructors agreed that the trial rubric was much better than the rubric that they had used

in other courses in the program. When asked about their agreement with the statement, “The quality of student writing in this course was better than previous courses,” 25 percent strongly agreed, 50 percent agreed, and 25 percent were neutral. This suggests that the rubric may have had some positive effect on student writing quality. Instructors were provided with a place to provide open comments about their experience with the rubric. Here, instructors pointed out several benefits of the rubric, including that it scaffolded better student writing, helped to curb grade inflation (associated with the previous use of administrative points), and produced a more accurate assessment of student performance.

Student Perceptions

Once the courses concluded, students were asked to provide feedback on their experience using the new rubric. Sixteen students out of 72 responded (response rate = 22 percent). Those who had taken at least one other course in the program were asked to rate their experience with this rubric compared to the previous one. Opinions were slightly negative, with 23 percent of students saying that the rubrics were somewhat better, 38 percent of students saying that it was about the same, 23 percent saying that it was somewhat worse, and 15 percent saying that it was much worse. Students also responded with their level of agreement to a series of statements about the rubrics. For analysis, Likert scores were centered on zero from -2 (strongly disagree) to 2 (strongly agree). Overall, students slightly disagreed with the statements: “I used the essay rubric in this course to help guide my writing process” (mean = $-.07$); “I found the essay rubric to be helpful in writing my essay” (mean = $-.07$); and “I was satisfied with the feedback that I received on my essay” (mean = $-.15$). They slightly agreed with the statement, “I referred to the rubric when writing my essay” (mean = $.15$). The most prominent theme that emerged from the open comments was confusion regarding the point values associated with the new rubric. Some students were confused by the dramatic reduction from 325 points to 50 points and suggested that this was not consistent with the amount of work involved in the essay task. They did not realize that the weight of the assignment in the scheme of assessment was still the same regardless of the number of points used to assess the essay.

Lessons Learned

Several lessons were learned from this first iteration of the rubric trial, from both the instructor and student perspectives. Instructors reacted positively to the rubric because they found it easier to use. They also opined that it was better at scaffolding student writing because of the new criteria for the introduction and the conclusion, but some expressed reservations about the weighting of those criteria. Although essay scores started out lower using the trial rubric there seemed to be an acclimation effect with continued use. This demonstrated that students learned to use the rubric to improve their performance. Students were mixed in their view of the rubric and seemed to be confused about the assignment weighting, which indicated a need for better communication from the project team. The most important lesson was that restricting the number of points in the rubric did not resolve challenges related to interrater reliability. This pointed toward the need for an additional mechanism to help instructors assess student performance more reliably and consistently.

Spring A 2023, 100-point Essay Rubric

Purpose and Structure

Several modifications were made to the essay rubric based on the feedback that was received in the first iteration. First, to address student concerns about the dramatic reduction in the number of points and the erroneous perception that essays carried less weight, the RWG increased the number of points in the rubric from 50 to 100 and improved the messaging about the scheme of assessment. Second, while the rubric categories remained the same, the RWG reduced the weighting attached to the introduction and conclusion and placed more emphasis on analysis and support. The new iteration of the trial essay rubric allocated points across the five assessment criteria as follows: introduction (5 points), analysis (35 points), support (35 points), conclusion (5 points), and communication (20 points). Third, to further scaffold student writing and guide instructor grading of the essays, the RWG developed extensive indicative content for each assignment. This indicative content was inspired by the mark schemes that senior readers write for College Board Advanced Placement exams in the United States and principal examin-

Table 4. Grade differences for Spring A 2023 trial

Course	Essay 1			Essay 2			Essay 3		
	Trial	Traditional	Diff	Trial	Traditional	Diff	Trial	Traditional	Diff
AP-5510	85.9% (n=34)	91.2% (n=17)	-5.3%*	90.2% (n= 34)	85.7% (n=17)	4.5%			
NS-5510	89.1% (n=29)	90.4% (n=43)	-1.3%	89.0% (n=29)	91.6% (n=43)	-2.6%**	90.2% (n=29)	90.6% (m=43)	-.4%
WS-5510	88.8% (n=30)	90.9% (n=32)	-2.1%*	89.4% (n= 30)	92.2% (n=32)	-2.8%**	90.5% (n=30)	91.7% (n=32)	-1.2%

** p < .01 * p < .05
Source: courtesy of the authors.

ers develop for General Certificate of Secondary Education and General Certificate of Education Advanced Level exams in the United Kingdom. The significance of these quality assurance documents is discussed in the section of this chapter titled “Some Key Lessons from the GCPME Rubric Trials.”

Grades

Grades were extracted from the courses using the same process used for the first iteration. Aggregate scores and the results of the independent samples t-test are displayed in table 4. Like the Summer B trial, scores from the trial rubric sections were lower than those using the traditional rubrics. However, the differences were not quite as large as those in the previous iteration.

Instructor Feedback

Instructors provided feedback on the rubric and the indicative content through a survey. Five instructors (four of whom were new to this iteration) responded to the survey. Again, they found the rubric easy to use (80 percent strongly agree, 20 percent agree), that it helped to focus on the most important parts of student writing (20 percent strongly agree, 80 percent agree), that it was a good way to assess student work (80 percent strongly agree, 20 percent agree), and that it was helpful in guiding feedback on student writing (60 percent strongly agree, 40 percent agree). All

instructors agreed that the trial rubric was much better than the rubric that they had used in other courses in the program, but they were more tepid in their response compared to the first iteration. In this round, 50 percent cited the rubric as much better and 50 percent cited it as somewhat better, whereas in Summer B, all instructors rated it as much better. Also, instructor perceptions were slightly positive regarding the quality of student writing in the sections that used the trial rubrics. When asked about their agreement with the statement “The quality of student writing in this course was better than previous courses,” 80 percent agreed and 20 percent disagreed.

Instructors were also asked about the use of indicative content to support essay writing. Overall, the attitude toward indicative content was mixed. While all the instructors saw evidence that students were using the indicative content to support their essay writing, mean agreement to the statement “Indicative content improved the quality of my student’s writing” was 3 (on a scale from 1 to 5). When asked whether the indicative content helped instructors to assess student writing, the score skewed toward disagreement (with a 2.8). In the open comments section of the survey, instructors were similarly positive about the rubric while somewhat tepid about the indicative content. Several instructors commented that the rubric allowed for a “more accurate assessment of student learning” and supported better student writing, but they felt that the indicative content was challenging because it was confusing, made grading harder, and “provided another set of readings beyond the assigned readings for the week.”

Student Feedback

In this iteration, a total of 30 students provided feedback on the use of the rubric and indicative content via survey, with results showing some positive trends, with 68 percent of students rating the rubric as better (23 percent much better and 45 percent somewhat better) than previous rubrics. This was an increase from the Summer B trial, where only 23 percent of students rated the rubric as somewhat better. Additionally, more students indicated that they found the rubric to be helpful when writing their essay (74 percent in Spring A versus 70 percent in Summer B), and

that they were satisfied with the feedback that they received on their essay (89 percent in Spring A versus 76 percent in Summer B).

Regarding the indicative content, nearly all (97 percent) of the students responded that they had looked at this resource and most (76 percent) understood its purpose. Unfortunately, no survey data was collected regarding whether students found the indicative content helpful in writing their essays. However, a rudimentary analysis of the free response comments from students identified several themes regarding their use of the rubric and the indicative content. On the positive side, students suggested they were helpful and saved time because they provided extensive advice and guidance about analytical themes, relevant examples, essay structure, writing style, and useful sources. However, other student responses also indicated several criticisms. Some suggested that the indicative content felt bolted onto the course and needed to be better integrated into the lessons and readings. This would avoid conflict with other sources of information and guidance. There was a perception that the indicative content had additional citation requirements beyond those prescribed in the GCPME style guide.²⁷ Some students suggested that rubric weightings need to be adjusted with less emphasis on communication, introduction, and conclusion.

Lessons Learned

This iteration of the essay rubric trial indicated incremental gains in rubric usability for instructors. Instructors were positive toward the rubric, suggesting that it was easy to use and provided a more accurate assessment of student learning. These comments may explain the continued difference in grades that was observed when comparing the trial and traditional sections. Instructor reaction to the indicative content was mixed, with some suggesting that it did not significantly improve student writing or help with grading. Student reactions were similar, and some found it confusing and poorly integrated into the course. In response to this feedback, the RWG worked with learning architects and instructional designers to integrate the indicative content directly into the assignments in the

²⁷ *The eSchool Style Guide* (Maxwell AFB, AL: Air University, 2019).

Canvas learning management system. This has enhanced the scaffolding of assignments in the OLMP and IDE programs.

At this point, the RWG was confident that it had achieved proof of concept with the 100-point essay rubric. It recommended rapid testing of a 100-point rubric for online discussions, with the objective of implementing both rubrics simultaneously in fall 2023. However, by this stage, it was becoming apparent that OBME requirements were pushing GCPME rubric development in a different direction. OBME assessment focused on program-level rubrics designed to assess program effectiveness: “Program-level rubrics do not grade student artifacts; rather, they assess them for PLO [Program Level Outcomes] attainment. These rubrics can be very effective in program-level assessment, but they require development at the program level and a second assessment process involving the aggregation of student artifact data.”²⁸ The OBME rubrics are more complex than the GCPME trial rubrics, with more criteria and variable point levels. This militated against the RWG objective that rubrics should be easier for instructors to use. The RWG pushed ahead with discussion rubric trials because it assumed that discussion rubrics would not be subject to OBME requirements. We later discovered that this assumption was mistaken. However, the OBME rubrics have incorporated some of the beneficial features of the trial rubrics, for both essays and discussions.

Results: March 2023, 100-point Discussion Rubric

Rubric Purpose and Structure

The original GCPME discussion rubric used 350 points distributed across six assessment criteria: analysis (100 points), course concepts and support (100 points), writing style (50 points), writing mechanics (50 points), interaction (25 points), and administrative (25 points). The RWG had several goals in developing a trial discussion rubric. First, it aimed to simplify the rubric by reducing the number of assessment criteria, which would make discussions easier to grade. The RWG was cognizant that instructors grade more discussions than essays, so fewer criteria

²⁸ *Outcomes-Based Military Education Procedures for Officer Professional Military Education*, D-B-1–D-B-2.

would facilitate easier grading. Second, the RWG wanted to incentivize discussion activity. Its objective was to shift the pattern of student engagement away from the required initial post no later than each Thursday and follow-on posts no later than each Sunday toward a pattern of earlier and more regular engagement throughout the week. The mechanism for this was to increase the weighting for the interaction category from just 7 percent of the score in the traditional rubric, to 50 percent of the score in the trial rubric. Third, as with the trial essay rubric, the elimination of administrative points was intended to address distortions in student scores. Fourth, the use of a unified communication criterion reduced the previous weighting of 31 percent attached to writing style and writing mechanics. The RWG agreed to the following distribution of points for the trial rubric: course concepts and support (40 points), interaction (50 points), and communication (10 points).

Grades

The same extraction and data analysis methods were used to process grades for this rubric trial and grade data is displayed in table 5. Like the results seen in the essay rubric trials, the scores from the trial discussion rubric sections were lower than those using the traditional rubrics. However, as with previous trials, differences in student scores abated over time.

Instructor Feedback

The RWG received feedback from all 13 instructors that participated in the trial. The survey revealed that 69 percent of instructors felt that the rubric was better (23 percent much better and 46 percent somewhat better) than previous discussion rubrics that they had used in GCPME courses, while 31 percent said that the rubric was worse (8 percent somewhat worse and 23 percent much worse). Instructors were then asked about their agreement level with a series of statements. Overall, all but one of the statements leaned toward agreement (on a scale of 1 to 6). In order of mean strength, they were:

- 4.38 I found the discussion rubric in this course easy to use

Table 5. Grade differences for March 2023 trial

Course	Discussion 1			Discussion 2			Discussion 3		
	Trial	Traditional	Diff	Trial	Traditional	Diff	Trial	Traditional	Diff
ISS-602A	86.2% (n=84)	92.8% (n=173)	-6.6%**	86.7% (n=84)	92.7% (n=173)	-6.0%**			
LDR-602A	82.1% (n=68)	89.5% (n=210)	-7.4%**	88.1% (n=68)	92.0% (n=210)	-4.1%**			
AIR-603A	85.5% (n=65)	91.2% (n=193)	-5.7%**	87.0% (n=65)	92.7% (n=193)	-5.7%**	89.5% (n=65)	93.0% (n=193)	-3.5%**

** p < .01 * p < .05

Source: courtesy of the author.

- 3.93 Allowed me to differentiate more easily between levels of student performance
- 3.85 Allowed me to grade more quickly than previous rubrics I've used
- 3.77 My students understood the requirements in the rubric
- 3.46 Facilitated a positive shift in peer-to-peer engagement
- 3.46 Facilitated a positive shift in student-to-instructor engagement
- 2.85 The level of student engagement in the discussion activities was better in this class than in previous GCPME courses I've taught

A brief analysis of the open response section of the survey shows instructors were divided about the effects of the discussion rubric. Most instructors thought the rubric was simpler and clearer, allowing for faster grading. However, some thought it was slower to grade because of the need to focus on the quantity of student comments. Some instructors noted that although scores decreased, the rubric produced accurate assessment that did not inflate grades. Most instructors saw shifts in the pattern of student engagement, but some were doubtful about this. Many

of the instructors provided welcome feedback for improving the rubric, which included suggestions to restore the critical thinking criterion, decrease the weighting attached to engagement, and provide a mechanism for addressing administrative issues such as the submission of late work.

Student Feedback

Thirty-six students responded to the survey asking for feedback on the discussion rubric. When asked about how the trial discussion rubric compared to previous rubrics they had used, perceptions skewed toward the negative (mean = 2.86). Only 28 percent of students felt that the rubric was better (14 percent much better and 14 percent somewhat better), 21 percent of students thought it was about the same, and 50 percent felt that it was worse (43 percent somewhat worse and 7 percent much worse). Students indicated that they referred to the rubric when writing [their] discussion responses (mean agreement = 4.47). They agreed with the statement that they found the rubric to be helpful in completing [their] responses (mean agreement = 3.88) and that the discussion rubric encouraged [them] to be more active in the discussion activity than [they] might normally be (mean agreement = 3.75). The analysis of open comments revealed that while a few students thought the rubric was helpful and encouraged them to be more active than they would otherwise be, most did not like the fact that the trial rubric required additional engagement. Students were candid in expressing their discontent about this because it put additional pressures on their already busy lives.

Lessons Learned

The central lesson learned from this rubric trial concerned student engagement. The RWG recognized that it needed to strike a balance between what was desirable and what was possible. Research has demonstrated consistently that student engagement is undoubtedly correlated with better educational outcomes at all educational levels. Yet, in the case of PME, students have competing demands that militate against greater engagement. Nevertheless, GCPME did not abandon this objective. Discussion rubrics for OBME assessment now include a “Community” criterion, which encourages and rewards students for exceeding minimum

Table 6. Grade differences for June 2023 trial

Course	Essay 1			Essay 2			Essay 3		
	Trial	Traditional	Diff	Trial	Traditional	Diff	Trial	Traditional	Diff
ISS-602A	90.8% (n=87)	90.0% (n=126)	0.8%	89.5% (n=87)	89.4% (n=126)	0.1%	90.7% (n=87)	89.9% (n=126)	0.8%
LDR-602A	81.3% (n= 61)	90.2% (n=155)	-8.9%**	84.8% (n= 61)	93.2% (n=155)	-8.4%**			
AIR-603A	90.2% (n=75)	91.2% (n=161)	-1.0%	91.0% (n=75)	91.8% (n=161)	-0.8%	91.8% (n=75)	92.9% (n=161)	-1.1%

** p < .01

Source: courtesy of the author.

standards, but does not prejudice achievement of satisfactory scores if students are unwilling or unable to meet elevated expectations for engagement.

Results: June 2023, 100-point Discussion Rubric

Rubric Purpose and Structure

To address feedback received in the March 2023 discussion rubric trial, the RWG restored the critical thinking criterion, reduced the weighting for interaction, and rebalanced the point distribution as follows: critical thinking (30 points), course concepts and support (30 points), interaction (30 points), and communication (10 points). The RWG also wrote a new administrative policy for grading work that was submitted late or failed to meet other expectations short of academic integrity violations.

Grades

The same extraction and data analysis methods were used to process grades for this rubric trial. The grade data is displayed in table 6. Interestingly, compared with iteration three, there were no appreciable differences in grades between the trial and traditional sections in ISS-602A and AIR-603A, but there was a significant difference in LDR-602A. There is no additional data to explain this difference in outcomes. However, LDR-602A is the first facilitated course that students take in the IDE program, so they may be adapting to program expectations in their PME.

Instructor Feedback

Thirteen instructors provided feedback on the trial rubric. Overall, instructors were nearly unanimous in the preference of this rubric over previous discussion rubrics, with 92 percent saying that it was better (54 percent much better and 38 percent somewhat better). Instructors were asked to indicate their level of agreement with a series of statements, and there was a significant shift in mean levels of agreement:

- I found the discussion rubric in this course easy to use—up 1.08 from 4.38 to 5.46 (1.08)
- Allowed me to grade more quickly than previous rubrics I've used—up 1.3 from 3.85 to 5.15
- Allowed me to differentiate more easily between levels of student performance—up .99 from 3.93 to 4.92
- Facilitated a positive shift in peer-to-peer engagement—up .69 from 3.46 to 4.15
- Facilitated a positive shift in student-to-instructor engagement—up .54 from 3.46 to 4.00
- The level of student engagement in the discussion activities was better in this class than in previous GCPME courses that I've taught—up 1.07 from 2.85 to 3.92

The open comments uncovered similar themes to those found in the first iteration of the discussion rubric trial. Instructors thought the rubric made grading easier and resulted in scores that better reflected what instructors thought students deserved for their work. Instructors liked that the rubric incentivized additional interaction, but some instructors noted that students did not change their behavior following feedback and seemed to be comfortable with meeting only the minimum requirements. Some instructors also suggested that the weighting of criteria could be further adjusted and that there could be better disambiguation between what was minimally required and what was maximally encouraged.

Student Feedback

As with the previous iteration, 36 students responded to the request for feedback. In analyzing their responses, the RWG did not see a cor-

relation between student and instructor views of the trial rubric. When asked to compare this version of the discussion rubric with versions that they had used in previous courses, only 25 percent said that it was better (5 percent much better and 20 percent somewhat better), which was lower than the 28 percent from the March iteration, while 40 percent said that it was about the same and 35 percent said that it was worse (25 percent somewhat worse and 10 percent much worse). There were either not or only marginal increases in the level of agreement with the statements: “I found the rubric to be helpful in completing my responses” (mean agreement = 3.88, no increase over March); and “The discussion rubric encouraged me to be more active in the discussion activity than I might normally be” (mean agreement = 3.91, increase of 0.16). Most of the students’ open comments were critical toward the new rubric, emphasizing a dislike of the additional interaction requirements. The RWG anticipated these results because the intent to promote greater student engagement was the same between the March and June discussion rubric trials.

Some Key Lessons from the GCPME Rubric Trials

The GCPME rubric trials demonstrate some key takeaways that may guide other institutions considering changes to their own rubrics. The most important lesson is that sound rubrics are a necessary but not a sufficient condition for improving teaching, learning, and assessment. In other words, rubrics will not on their own solve key problems emanating from the inputs to and outputs of the teaching, learning, and assessment process. Looking at inputs, for example, one objective of the trial discussion rubrics was to encourage more regular student engagement in formative assessment. For graduate students especially, engagement in online discussions promotes a better social learning experience, provides more opportunities for instructors to offer feedback, and is closely correlated with deeper learning and improved performance in summative assessment.²⁹ However, the culture of PME—at least in the mandatory programs—is

²⁹ See Min Young Doo and Jungwon Kim, “The Relationship between Learning Engagement and Learning Outcomes in Online Learning in Higher Education: A Meta-analysis Study,” *Distance Education* 45, no. 1 (2024): 60–82, <https://doi.org/10.1080/01587919.2024.2303484>.

highly resistant to challenging students to exceed minimum expectations for engagement.³⁰ Many students assert that because they have demanding military careers and family commitments, it is unreasonable to ask them to demonstrate greater presence in discussions. Their extrinsic barriers to learning therefore become deeply internalized, and PME is perceived as a check-box exercise, the only purpose of which is to facilitate promotion. Daniel Hughes captured this memorably (and perhaps notoriously) in his discussion of residential students at the Air War College, half of whom he suggested “have no interest in attendance and would not attend if they could otherwise be certain of advancement to colonel.” The consequence, Hughes observed, is that many students attend classes “without enthusiasm, or any intellectual curiosity,” demonstrating attitudes “similar to those of disinterested undergraduates.”³¹ One way to meet this challenge is to anchor assessment tasks to DOD directives, which helps to establish the relevance of what students are being asked to do by connecting it explicitly to their mission.³² This approach has proven very successful in one of the short IDE courses at GCPME: ISS-602A, “Making Decisions: Strategy and War,” in which students apply theoretical approaches to strategic competition with China. Student satisfaction rates on that course consistently exceed 95 percent and the end-of-course comments frequently highlight the relevance of authentic assessment methods as a strength of the course. This suggests that affective barriers can be overcome if sufficient attention is given to assessment design and purpose.

Turning to the outputs of the teaching, learning, and assessment process, it is evident that rubrics will not provide a silver bullet for instructor-related issues including feedback deficiency, grade inflation, and lack of interrater reliability. A well-crafted rubric will offer instructors

³⁰ Daniel Hughes, “Professors in the Colonel’s World,” in *Military Culture and Education*, ed. Douglas Higbee (Burlington, VT: Ashgate Publishing, 2010), 152.

³¹ Hughes, “Professors in the Colonel’s World.”

³² For example, *Developing Today’s Joint Officers for Tomorrow’s Ways of War*; Gen Lloyd J. Austin III, U.S. Secretary of Defense, memorandum, “Professional Military Education Learning Outcomes on the People’s Republic of China,” 8 July 2022; and “Appendix II: Professional Military Education China-Focused Learning Outcomes,” *Military Officers: DOD Can Enhance Promotion and Education Guidance for Addressing Indo-Pacific Region Needs* (Washington, DC: Government Accountability Office, 2023), 28–29.

clear opportunities to provide sufficient feedback using the language in the rubric cells in combination with the assessment instructions, required and recommended resources, and their own subject matter expertise. However, just as levels of student engagement and performance vary considerably, the same is also true of instructors. Students complain that feedback is often inadequate for various reasons, including lack of consistency, specificity, and clarity.³³ This demands vigilance and diplomacy from course directors who are tasked with managing dozens of contract instructors, many of whom have other jobs and do not always approach their adjunct responsibilities commensurate with the seriousness of the PME mission. One method that can be used to model what PME institutions are looking for from their contract instructors is to distribute anonymized exemplars that demonstrate quality feedback. These exemplars must use the rubric language explicitly and annotate the text to indicate how students do not satisfy, meet, or exceed assessment standards.

A well-crafted rubric will allow instructors to differentiate effectively between different levels of student performance, but this does not automatically guard against grade inflation or grade clustering, nor does it guarantee improvements in interrater reliability. According to Hughes, in 2009, one of the departments at the Air War College awarded grades of A and A- to 97 percent of students in a core course.³⁴ At GCPME, some part-time instructors tend to award grades that are, in the aggregate, rather higher than those awarded by full-time faculty. This is a longstanding and well-known phenomenon not restricted to the PME community. It has become more common because of the decline in the number of permanent, full-time faculty and the greater reliance on temporary, part-time faculty in U.S. higher education institutions. The commercialized or corporatized institutions of higher education have produced several deleterious effects, including the rise of a consumer mentality among students.³⁵ Adjunct instructors have responded by awarding higher grades because they

³³ Wells, "Building Trust and Success Through Dialogic Feedback in Joint Professional Military Education," 58.

³⁴ Hughes, "Professors in the Colonel's World," 154.

³⁵ Fredrik DeBoer, "Why We Should Fear University, Inc.," *New York Times Magazine*, 13 September 2015.

wish to garner favorable student evaluations, which they perceive as the currency with which they can purchase future short-term contracts in a highly competitive employment market.³⁶ Institutions have been complicit in this because they wish to retain students, improve graduation rates, and protect their financial position. This creates competing pressures between faculty integrity and institutional success.³⁷ A recent study indicated that adjunct faculty did not consider themselves or the institutions in which they taught to be contributors to the problem of grade inflation. Interestingly, none of the faculty surveyed in that study had received any training about grade inflation.³⁸ This draws attention to the significance of faculty training to ensure fair and consistent assessment practices.

From an early stage in the GCPME rubric trials, it became clear that the introduction of new rubrics is only the first step in a process, and ongoing faculty training will be necessary to ensure the proper application of rubrics. This can be achieved through rubric norming, which is intended to promote interrater reliability, where “norming is now recognized as an important aspect of rubric use and is crucial to the reliability and validity of a rubric. Without such a process, deployment of a rubric may be a waste of time, or severely limit its effectiveness.”³⁹ Rubric norming is used routinely in K-12 education. In the United Kingdom, for example, academic assessment leading to the award of educational qualifications such as the General Certificate of Secondary Education and General Certificate of Education Advanced Level is governed by rigorous quality assurance arrangements. The Office of Qualifications and Examinations Regulation mandates that examining boards must use a hierarchical system of principal examiners and senior examiners, who supervise thousands of examiners. Standardization (rubric norming) sessions are held to ensure that

³⁶ Brenda S. Sonner, “A Is for ‘Adjunct’: Examining Grade Inflation in Higher Education,” *Journal of Education for Business* 76, no. 1 (2000): 5–8, <https://doi.org/10.1080/08832320009599042>.

³⁷ David E. Smith and Steven Fleisher, “The Implications of Grade Inflation: Faculty Integrity versus the Pressure to Succeed,” *Journal of Research in Innovative Teaching* 4, no. 1 (2011): 32–38.

³⁸ Billy C. Johnson and Danny E. Malone Jr., “Not Me, Not Here: Adjunct Faculty Perceptions of Grade Inflation at U.S. Colleges and Universities,” *American Journal of Qualitative Research* 7, no. 2 (2023): 147–62, <https://doi.org/10.29333/ajqr/13137>.

³⁹ Schoepp, Danaher, and Kranov, “An Effective Rubric Norming Process,” 4.

examiners understand the mark scheme (rubric) and thereafter senior examiners sample the scripts graded by examiners to ensure that the mark schemes are being applied consistently and fairly.⁴⁰ In the United States, the College Board uses a similar system of quality assurance to standardize assessment for Advanced Placement exams. The chief reader for each exam, who is always a faculty member at an institution of higher education, develops rubrics and selects reading leadership and readers who score free-response questions. Readers receive thorough training to ensure that they understand the rubrics and their work is sampled throughout the exam reading, which takes place over several days, for fairness and consistency.⁴¹

An analogous system in PME institutions would see department chairs serving a coordinating role as principal examiner and course directors acting as chief examiners to monitor the examiners (contract instructors). Adoption of such a definitive hierarchical structure would be very much in keeping with military culture and would go some way toward meeting concerns expressed in a recent Rand workshop on PME that there is a “need for experts who understand the curriculum to build creative academic methodologies and authentic assessments and write rubrics to guide talent management.”⁴² The phrase “guide talent management” implies a process of continuous improvement, but this represents a significant challenge for an institution like GCPME in connection with faculty training on assessment. Most of its instructors are adjuncts and the time and resources to support professional development for hundreds of instructors is lacking. Further, the process of professional development is mediated by outsourcing. Contract instructors are drawn from a community of retired military officers and civilian educators, but they do not report directly to the course directors who are responsible for the curricu-

⁴⁰ Martin Johnson, “Articulation Work: Insights into Examiners’ Expertise from Their Remote Feedback Interactions,” *Communication & Language at Work* 4, no. 4 (2015): 30–31, <https://doi.org/10.7146/claw.vli4.20771>.

⁴¹ “How AP Develops Courses and Exams,” Ap Central, College Board, accessed 25 September 2024.

⁴² Emily Ellinger and Mark N. Posard, *Imagining the Future of Professional Military Education in the United States: Results from a Virtual Workshop* (Santa Monica, CA: Rand, 2023), 16, <https://doi.org/10.7249/CFA2148-1>.

lum and assessment. The instructors are subject matter experts, but they are not experts in assessment. This should not be surprising, as very few higher education faculty receive specific and ongoing training in assessment and the number of PME faculty who have benefitted from such professional development is even smaller. Clearly, then, a system for rubric norming and faculty training should be a priority if the PME community is to claim that its assessment practices meet the exacting standards laid down by OBME requirements for validity, reliability, and calibration.⁴³

Conclusion

This chapter has reported on an educational quality improvement study, which implemented principles and processes from DBR and the literature on rubric development to iterate and refine rubrics used to assess online discussions and essay writing at GCPME. The design and use of rubrics has attracted more attention in higher education recently because accreditation bodies have compelled a sharper focus on effective assessment. However, there is little evidence that the PME community had considered the significance of this until the Directorate for Joint Force Development thrust the issue into focus as part of the transition to OBME. PME institutions should regard this as only a first step in the systematic consideration of rubrics. Although OBME rubrics will facilitate the collection of program-level data that can be reported upward, the validity of the data will be brought into question unless there is a robust mechanism for quality assurance in PME assessment practices. Ensuring interrater reliability through regular calibration is likely to be a major challenge for PME institutions operating at scale with hundreds of instructors and thousands of students. However, the qualification authorities that govern public examination systems in the United States and the United Kingdom offer a model that can be adapted for use in the PME community. PME institutions should also look beyond the use of rubrics for reporting valid and reliable assessment data. Best practice emphasizes that rubrics should be used to scaffold teaching and learning. When rubrics are used to support

⁴³ *Outcomes-Based Military Education Procedures for Officer Professional Military Education*, Enclosure D, Appendix B, “Rubric Guidelines and Examples.”

transparent assessment, students can improve their subsequent performance by using effective feedback: “Such transparency and feedback increase student self-confidence, self-efficacy, self-awareness, and self-determination.”⁴⁴ The GCPME rubric trials were predicated first and foremost on this aim. Although eventually overtaken by OBME requirements, they demonstrated a sustained attempt to improve teaching and learning at GCPME and served as an excellent example of design-based research in the PME environment.

⁴⁴ Ragupathi and Lee, “Beyond Fairness and Consistency in Grading,” 91.



CHAPTER 10

Experiments and Testing

Evidence-Based Instructional Strategies
at the U.S. Naval Community College

*Charles B. Cushman Jr., PhD; Russell Evans;
Daniel Goff; Scott Hamm, PhD; and Thomas Robertson*

Introduction

Faculty at the U.S. Naval Community College (USNCC) have been testing several teaching delivery modes and numerous delivery techniques to improve its Naval Studies Certificate courses. Each experiment is assessed, with the resulting data driving decision making on future programming, course design, and delivery modes. From its initial inception, USNCC has drawn on expertise from the full Naval University System to design its courses and certificate programs and has used assessment, student survey, and faculty evaluation to judge the effectiveness of the courses and to suggest revisions.¹ At the same time, USNCC faculty test techniques for improving student retention and engagement with the material and conduct experiments with a variety of delivery modes, including specific courses and testing delivery to ships underway, to fulfill its mission to support educational opportunities for U.S. Navy/Marine Corps/Coast Guard enlisted personnel.

Following recommendations made in the *Education for Seapower* study (2018), the U.S. Department of the Navy (DON) established the USNCC

¹ MC2c Kashif Basharat, “Navy Renews Focus on Education: Service Announces University System,” *All Hands Magazine*, 12 February 2019.

in 2019 to provide naval-relevant associate degrees and certificates to enlisted members of the Navy, Marine Corps, and Coast Guard to enhance readiness, improve warfighting skills, and prepare students to take advantage of lifelong learning.²

The USNCC team developed a three-part program model that is taught in partnership with key civilian higher education institutions offering high-quality relevant programs. Each USNCC program consists of a naval studies certificate, which makes up part of a core curriculum focused on twenty-first-century intellectual skills, and a professional concentration leading to an embedded certificate; the three parts together (naval studies, core, and concentration) lead to an accredited associate degree.

USNCC faculty developed a five-course, 15-credit naval studies certificate to provide a baseline of *navalized* skills and knowledge to USNCC students, which transfers to the partner institutions as part of their core curricula. USNCC and the partner schools worked together to develop degree and certificate programs in a pilot program that ran until September 2023, when USNCC achieved its initial operating capacity, with 3,000 students enrolled in nine associate degree programs operated by six partner institutions.³

Because USNCC serves a globally deployed student body who pursue their classes after duty hours, USNCC (and its partners) offers all its courses in an online, asynchronous mode. As a new school seeking to maximize opportunities to reach its students, USNCC is committed to testing other means of teaching and course delivery. Building a learning commu-

² *Education for Seapower: Final Report, December 2018* (Washington, DC: Department of the Navy, 2018), 7.

³ *State of the College, 2023: Pilot II Report* (Quantico, VA: U.S. Naval Community College, 2023).

nity to serve this globally deployed population is fundamental to success.⁴ With this in mind, USNCC considered two key research questions that would guide initial course and program design, as well as underpin ongoing efforts to develop a robust faculty-student learning community:

- How can USNCC improve naval studies courses and the certificate program by using additional in-course techniques to increase student learning and engagement?
- How might USNCC expand its reach and help more prospective students study at USNCC via alternative delivery modes?

USNCC was founded as an online educational institution based on prior DON education programs, as explored in the *Education for Seapower* report; only online education can reach the sailors, Marines, and Coast Guard personnel deployed worldwide, and staffing demands prevent sending them to in-residence programs beyond their initial training.⁵ As a result, USNCC did not have to pivot to online courses when COVID-19 hit, as that was already the course delivery model.⁶ USNCC institution-

⁴ Florence Martin and Doris U. Bolliger, “Designing Online Learning in Higher Education,” in *Handbook of Open, Distance and Digital Education*, ed. Olaf Zawacki-Richter and Insung Jung (Singapore: Springer, 2023), https://doi.org/10.1007/978-981-19-2080-6_72; Jeffrey S. Lantitis, ed., *Active Learning in Political Science for a Post-Pandemic World: From Triage to Transformation* (Cham, Switzerland: Palgrave Macmillan, 2022); and Sally J. Baldwin, Yu-Hui Ching, and Norm Friesen, “Online Course Design and Development among College and University Instructors: An Analysis Using Grounded Theory,” *Online Learning Journal* 22, no. 2 (2018): 157–71, <http://dx.doi.org/10.24059/olj.v22i2.1212>.

⁵ *Education for Seapower*.

⁶ Richard Garrett et al., *CHLOE 7: Tracking Online Learning from Mainstream Acceptance to Universal Adoption: The Changing Landscape of Online Education* (Annapolis, MD: Quality Matters, 2022); Sebastian Kaempf and Carrie Finn, “Teaching International Relations through the Format of a Massive Open Online Course (MOOC),” *International Studies Perspectives* 22, no. 1 (2021): 1–24, <https://doi.org/10.1093/isp/ekz025>; Taiyi Sun, “Forced Experimentation: Teaching Civic Engagement Online amid COVID-19,” *PS: Political Science and Politics* 54, no. 1 (2020): 1176–79, <https://doi.org/10.1017/S1049096520001559>; Houman Harouni, “Unprepared Humanities: A Pedagogy (Forced) Online,” *Journal of Philosophy Education* 55, no. 4–5 (2021): 633–48, <https://doi.org/10.1111/1467-9752.12566>; and Barbara Oakley, Debra Poole, and MaryAnne Nestor, “Creating a Sticky MOOC,” *Online Learning Journal* 20, no. 1 (March 2016): 1–12, <https://doi.org/10.24059/olj.v20i1.731>.

al design features small sections averaging 25 students in asynchronous, online courses of eight weeks duration, offered in two fall and spring sessions and one summer session. Robust student coaching supports the students as they pursue their coursework; this is provided by both USNCC and its partner schools (see the next section for a full description of the partnership model).

USNCC faculty developed the five-course Naval Studies Certificate using identified best practices for online courses, including a comprehensive course outline and introduction to orient students, a standardized user interface and course design, well-designed readings and assignments, and real-world examples and situations that demonstrate the relevance of the coursework to the students' careers.⁷ The faculty designed the courses and sequence with USNCC leadership, expert practitioners from the Services, and subject matter experts at the other Naval University System schools: the Naval Academy, Marine Corps University, the Naval

⁷Peter Kilgour et al., "Threshold Concepts about Online Pedagogy for Novice Online Teachers in Higher Education," *Higher Education Research & Development* 38, no. 7 (2019): 1417–31, <https://doi.org/10.1080/07294360.2018.1450360>; Janet G. Auten and Margaret M. Twigg, "Teaching and Learning SoTL: Preparing Future Faculty in a Pedagogy Course," *Teaching & Learning Inquiry* 3, no. 1 (2015): 3–13; Lantis, *Active Learning in Political Science for a Post-Pandemic World*; Emily Lewis, "Best Practices for Improving the Quality of the Online Course Design and Learners Experience," *Journal of Continuing Higher Education* 69, no. 1 (2021): 61–70, <https://doi.org/10.1080/07377363.2020.1776558>; Nancy Chick, Lorelli Nowell, Bartłomiej A. Lenart, "The Scholarship of Teaching and Learning: A Scoping Review Protocol," *Teaching & Learning Inquiry* 7, no. 2 (2019), <http://dx.doi.org/10.20343/teachlearninqu.7.2.12>; Swapna Kumar et al., "Award-winning Faculty Online Teaching Practices: Elements of Award-winning Courses," *Online Learning Journal* 23, no. 4 (2019): 160–80, <https://doi.org/10.24059/olj.v23i4.2077>; Mansureh Kebritchi, Angie Lipschuetz, and Lilia Santiago, "Issues and Challenges for Teaching Successful Online Courses in Higher Education: A Literature Review," *Journal of Educational Technology Systems* 46, no. 1 (2017): 4–29, <https://doi.org/10.1177/0047239516661713>; Sandro Monteiro et al., "Course Design in E-learning and the Relationship with Attrition and Dropout: A Systematic Review" (paper delivered at Fourth International Instructional Technologies & Teacher Education Symposium, 6–8 October 2016, Firat University, Turkey); Tena B. Crews, Kelly Wilkinson, and Jason K. Neill, "Principles for Good Practice in Undergraduate Education: Effective Online Course Design to Assist Students' Success," *MERLOT Journal of Online Learning and Teaching* 11, no. 1 (March 2015): 87–103; Peter Felten, "Principles of Good Practice in SoTL," *Teaching & Learning Inquiry* 1, no. 1 (2013): 121–25, <https://doi.org/10.2979/teachlearninqu.1.1.121>; and Steven A. Meyers, "Using Transformative Pedagogy When Teaching Online," *College Teaching* 56, no. 4 (2008): 219–24, <https://doi.org/10.3200/CTCH.56.4.219-224>.

War College, and the Naval Postgraduate School. The faculty developed exercises, assignments, and discussion boards for each course based on research into effective online learning, suggesting that the foundation of effective online teaching is real-world connection between coursework and duties, first by subject matter relevant to the students' work responsibilities and then by use of examples and exercises with direct application to actual work circumstances.⁸ Building off that base, USNCC faculty also recognized the need to build community within courses because deeper student engagement increases the likelihood of course success; faculty achieve this through enhanced faculty presence in courses and through personalization, which seeks to meet the students where they are.⁹ All USNCC-taught courses emphasize faculty-student and student-student interactions, demonstrate relevance to their students real-world respon-

⁸ Martin and Bolliger, "Designing Online Learning in Higher Education"; Lantis, *Active Learning in Political Science for a Post-Pandemic World*; Baldwin, Ching, and Friesen, "Online Course Design and Development among College and University Instructors"; Kebritchi, Lipschuetz, and Santiago, "Issues and Challenges for Teaching Successful Online Courses in Higher Education"; Linda B. Nilson and Ludwika A. Goodson, *Online Teaching at Its Best: Merging Instructional Design with Teaching and Learning Research* (New York: John Wiley & Sons, 2017); Barbara Means et al., *Evaluation of Evidence-Based Practices in Online Learning: A Meta-Analysis and Review of Online Learning Studies* (Washington, DC: Office of Planning, Evaluation, and Policy Development, Policy and Program Studies Service, Department of Education, 2010); Meyers, "Using Transformative Pedagogy When Teaching Online"; and Scott D. Johnson and Steven R. Aragon, "An Instructional Strategy Framework for Online Learning Environments," *New Directions for Adult and Continuing Education*, no. 100 (Winter 2003): 31–43, <https://doi.org/10.1002/ace.117>.

⁹ Leanna Archambault, Heather Leary, and Kerry Rice, "Pillars of Online Pedagogy: A Framework for Teaching in Online Learning Environments," *Educational Psychologist* 57, no. 3 (2022): 178–91, <https://doi.org/10.1080/00461520.2022.2051513>; Lewis, "Best Practices for Improving the Quality of the Online Course Design and Learners Experience"; Sungwon Shin and Jongpil Cheon, "Assuring Student Satisfaction of Online Education: A Search for Core Course Design Elements," *International Journal on E-Learning* 18, no. 2 (2019): 147–64; John Steele, Rick Holbeck, and Jean Mandernach, "Defining Effective Online Pedagogy," *Journal of Instructional Research* 8, no. 2 (2019): 5–8; Martin and Bolliger, "Designing Online Learning in Higher Education"; Betul C. Czerkawski and Eugene W. Lyman III, "An Instructional Design Framework for Fostering Student Engagement in Online Learning Environments," *TechTrends* 60 (2016): 532–39, <https://doi.org/10.1007/s11528-016-0110-z>; and Monteiro et al., "Course Design in E-learning and the Relationship with Attrition and Dropout."

sibilities, and help students individualize/personalize their learning as they take their naval studies courses. Writing assignments across the courses all interlink, teaching students good research skills, organization, argument, and proper use of citations.¹⁰

From this theoretical basis, USNCC set up its initial courses and began amassing assessment data on their effectiveness, which pointed the faculty toward a series of experiments testing various in-course refinements, as well as alternative delivery modes.

Following a brief overview of the structure and approach of the naval studies program, this chapter examines how USNCC has sought to improve its in-course performance using experiments to enhance relevance, build community, and deepen student engagement with their coursework, as well as how USNCC has tested alternative delivery modes to increase its ability to serve a student body that is deployed worldwide. The chapter concludes with a summary of findings and recommendations for future experimentation as the basis for continuous improvement.

USNCC: Who Are We?

Naval Studies Program Design

The Navy launched USNCC in 2019 following a DON directive to establish the school. In the 2021 National Defense Authorization Act, Congress fully authorized USNCC to become a degree-granting institution once it achieves accreditation.¹¹ USNCC is currently pursuing accreditation through the Middle States Commission on Higher Education.

¹⁰ Alice S. Horning, “Critical Reading Skills: An Urgent Challenge,” *Inside Higher Ed*, 16 October 2022; John LaForest Phillips, “Not All Pain Is Gain: Lessons from Teaching Critical Thinking Online,” *PS: Political Science and Politics*, 54, no. 1 (2021): 178, <https://doi.org/10.1017/S1049096520001584>; Crews, Wilkinson, and Neill, “Principles for Good Practice in Undergraduate Education”; S. L. McMillan, “Bravo for Brevity: Using Short Paper Assignments in International Relations Classes,” *International Studies Perspectives* 15, no. 1 (2014): 109–20, <https://doi.org/10.1111/insp.12003>; and C. D. Coe, “Scaffolded Writing as a Tool for Critical Thinking: Teaching Beginning Students How to Write Arguments,” *Teaching Philosophy* 34, no. 1 (2011): 33–50, <https://doi.org/10.5840/teachphil20113413>.

¹¹ S.4049, William M. Thornberry National Defense Authorization Act for Fiscal Year 2021, 116th Cong. (23 July 2020).

As of this writing, USNCC enrolls more than 3,000 students in its programs and is envisioned in Navy planning documents to serve as many as 25,000 students by the time it reaches full operating capacity by or around 2028.¹²

The mission of USNCC is to “advance warfighting advantage and enhance operational readiness by providing world-class, naval-relevant education to a globally deployed force.”¹³ The college’s vision statement adds an important aspect to its approach:

The United States Naval Community College will develop warfighters across the force who are committed to lifelong learning, are steeped in naval traditions and values, have sound leadership qualities and ethical decision-making ability, possess improved critical thinking and technical skills, and have a deep understanding of the complex maritime environment in which they operate.¹⁴

In pursuit of this mission and vision, USNCC provides naval-relevant associate degrees and affiliated certificates to enlisted servicemembers in the Navy, Marine Corps, and Coast Guard. Initial program development ideas come from DON guidance, and USNCC refines them in close coordination with advisory groups drawn from senior enlisted leaders in the ratings and military occupational specialties served by the specific academic programs in the three naval Services. Following partner selection through the standard federal contracting process, USNCC develops the program with the partner, and USNCC faculty briefs the advisory groups on a regular schedule to ensure that the program as developed accurately captures Service needs.

USNCC solicits partner institutions with demonstrated expertise in the relevant fields, along with reputations for serving military students, and with the capacity to offer their programs in a fully online, asynchronous mode. USNCC offers its degrees through a consortium in partner-

¹² Shanessa Jackson, “USNCC Command Senior Enlisted Leader Jordan Rosado Speaks,” USNCC Newsroom, 2 February 2024.

¹³ *State of the College*, 2023, 2.

¹⁴ *State of the College*, 2023, 5.

ship with leading higher education institutions. USNCC degree programs currently consist of these subjects and partners:¹⁵

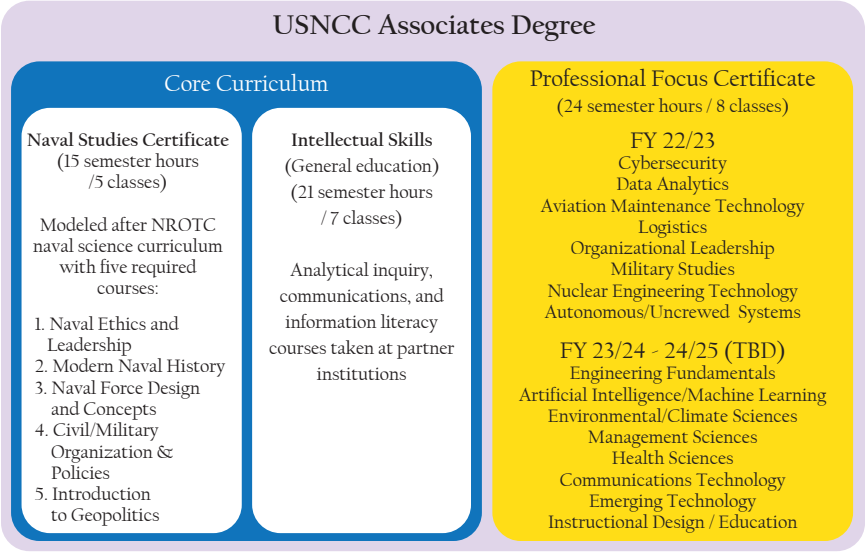
- Nuclear engineering technology: Alexandria Technical and Community College
- Aviation maintenance technology: Embry-Riddle Aeronautical University (ERAU)
- Uncrewed systems: ERAU
- Logistics (maritime): Arizona State University (ASU)
- Logistics: Tidewater Community College
- Military studies: ASU
- Organizational leadership: ASU
- Cybersecurity and information assurance: Western Governors University (WGU)
- Cybersecurity: Northern Virginia Community College
- General studies with computer studies specialization: University of Maryland Global Campus
- Data analytics: WGU

The program model used by USNCC is a three-part, stackable credential: the core curriculum focused on developing twenty-first-century skills offered at the partner school, the Naval Studies Certificate, and the professional concentration, which is also packaged as a stand-alone certificate. Figure 1 illustrates the stackable nature of the associate degree model that USNCC has constructed. All USNCC's consortium partners offer transfer credit using the students' Joint service transcripts, and all are recognized as schools that support military students. In addition, each program includes four-year pathways that treat the associate degree earned while at USNCC as fully transferable.

USNCC's unique offering is the Naval Studies Certificate, a five-course, 15 credit-hour sequence designed to deepen students' professionalism and dedication to their roles as servicemembers and to more fully connect their educational programs to their real-world duties. The program's learning outcomes lay out what USNCC seeks to teach in its naval

¹⁵ *State of the College*, 2023, 5.

Figure 1. Stackable design of the USNCC degree model



Source: courtesy of the authors, adapted by MCUP.

studies program; the first outcome captures the learning the students will do in their concentration field, while the rest focus on the courses in the partner schools' core curricula and the naval studies courses (sidebar). USNCC assesses all courses on a three-year cycle and works with partner schools to collect the relevant assessment data from them annually.

The naval studies courses follow a logic of expanding viewpoint: the first course, Naval Ethics and Leadership, focuses the individual students on their own roles as servicemembers and rising leaders with their Services, and the rest of the course sequence builds out from there to the role of their respective Service in the mission of the U.S. Department of Defense (DOD), the role of the DOD in the whole U.S. government, and the role of the United States on the global stage. Each course is eight weeks in length and includes numerous assessment and graded exercises that allow faculty to monitor progress on student learning. Each course is formatted using a standardized USNCC template, developed by USNCC faculty with an outside instructional design consultancy, so the courses all have the

Naval Studies Outcomes

1. Students shall demonstrate proficiency in one concentration area that has relevance to the Navy, Marine Corps, and Coast Guard Services.
2. Students shall demonstrate breadth of knowledge and the ability to make connections across a range of disciplines.
3. Students shall demonstrate twenty-first-century skills inclusive of critical thinking, analytical reasoning, information literacy, data fluency, and communicative fluency, inclusive of writing and verbal communications.
4. Students shall demonstrate learning, strong ethics, and judgment (individually or in teams) by addressing a problem using scholarly inquiry professionally or in other settings outside the classroom.
5. Students shall understand the context of the naval Services through the lenses of history, civil-military relations, and current geopolitical trends.

same look and feel, and student navigation through them is the same for the entire five-course sequence. Students develop research and writing skills through the sequence by completing increasingly complicated writing projects, culminating in a long paper and group presentations in NAV 105 (Introduction to the Geopolitical Environment).

Having devised a program and established courses, a template for course construction, and the asynchronous approach to delivering the Naval Studies Certificate, the USNCC Office of Institutional Research, Assessment, and Effectiveness worked with the faculty to design a program- and course-level assessment to track the effectiveness of the courses, program, and assignments. Guided by the findings of these efforts, the faculty began to develop and test other in-course approaches as potentially better tools for teaching; faculty also began testing alternative delivery modes to the asynchronous model in use.

In-Course Experiments

As USNCC began teaching its naval studies courses in 2023, the faculty conducted extensive assessment and student surveying in each iteration to gauge the effectiveness and relevance of the courses. This section reports on five experiments conducted by USNCC in the past two years. Prior to launching its courses, USNCC ran a joint test of the NAV 101 course with one of its partners to validate the content and delivery mode. Preliminary evaluation of student performance suggested that students who had already either taken English composition/EN 101 or had tested out of the requirement at one of USNCC's partner schools tended to perform better in naval studies courses than their peers who did not have that course in their background. This led the faculty to institute a new policy requiring satisfaction of the English requirement before enrolling in naval studies courses. This caused students to bounce back and forth between USNCC courses and

Naval Studies Courses

NAV 101 Naval Ethics & Leadership: Provides students with the foundational knowledge to assist them in understanding their role as a member of the U.S. military including their promise to “support and defend” the Constitution as well as be persons of exemplary character, worthy of moral praise.

NAV 102 Modern Naval History: Presents an overview of modern naval history, with an emphasis on understanding the roles that the naval Services (Marine Corps, Navy, Coast Guard) have played and how they relate to the relevant period of history and conflicts therein.

NAV 103 Naval Force Design: Instructs USNCC students in the organization of naval forces and fundamental concepts of naval operations and prepares students to recognize how naval power elements can be applied across strategic, operational, and tactical levels of warfare.

NAV 104 Civilian/Military Relations, Organization, and American Government: Provides students with an overarching understanding of the organization of the American political system and the role the military plays in that system.

NAV 105 Introduction to the Geopolitical Environment: Assists students in understanding the role of the United States in the international system, as well as the challenges it faces in the twenty-first century.

partner courses; USNCC faculty tested sending students through the Naval Studies Certificate courses in sequence before attending partner courses. USNCC also began testing the use of simulations and wargaming in select courses, given the importance of both in naval education, especially at the officer level. USNCC faculty also improved the relevance of the naval studies courses by enhancing the amount of material focused on the Coast Guard's roles and mission in maritime policy.

Testing NAV 101

USNCC leadership worked closely with other Naval University System schools (particularly the Naval Academy at Annapolis, Maryland, and the Naval Postgraduate School at Monterey, California) to produce a preliminary design of the five naval studies courses that would make up the certificate. With these designs in hand, USNCC worked with five partner schools in the Pilot I phase of operations (January 2021–January 2022) to offer NAV 101 at a partner school. This experiment allowed USNCC to: (1) assess the naval studies program design, (2) test the operation of institutional partnerships, and (3) to collect initial data on both student outcomes and relevance to the Services.

Preliminary data collected by USNCC supported the approach to course design and the delivery mode. The student surveys and faculty evaluation of NAV 101 indicated that the course was too long: it was more like a six-credit course than the desired three-credit course. This finding led the USNCC faculty to revise the course and align its content more accurately to the planned three-credit scale, which is the version being offered at USNCC now. The experiment also validated the consortium/partnership approach of USNCC and gave the staff valuable insight into how much time and effort would be required to bring on new partners when USNCC began to offer its initial associate degree programs in Pilot II (January 2022–September 2023).

English Readiness

During the first pilot, the success of students who self-identified as English-ready, either having taken a college-level English course or possessing high scores in English placement tests, were demonstrably better

than the level of success for those who did not. Of 185 students who did not report English-readiness, more than 10 percent earned grades of C- or lower, whereas less than 4 percent of 1,820 English-ready students earned those grades. This led to a decision to have students who did not already have a college English course begin with the partner institution and take an English course before beginning any classes in their degree focus. After examining attrition statistics most of the way through Pilot II, it became clear that although students performed well in NAV 101, the heavier writing requirements of NAV 102, including an 8- to 10-page research paper, seemed to overwhelm the students and resulted in significantly higher attrition. The naval studies courses are writing-intensive by design to promote critical thinking and increase communication skills, both of which can best be measured in an asynchronous course through evaluating argumentation and analysis in writing. With term lengths of only eight weeks, there is limited time for faculty to assist students whose writing skills are lacking. Due to the significant number of students who struggled with the writing, and its observed tie to attrition and predictor of success, English will become the first intellectual skill that USNCC will “navalize” and offer outside its Naval Studies Certificate. The English 101-like course will focus on arming students with the skills needed to write the reflective essays and research papers of the curriculum so that they can spend more time focusing on critical thinking, analysis, and communicating their thoughts. Both a one- and three-credit option will be offered to allow for differences in student needs. Those who have had a college English class before but need a refresher will be offered a one-credit short-duration course, while those who have not will be offered a three-credit class.

Cohorting Naval Studies

The purpose of the ongoing 3A pilot is to assess the relevancy of the Naval Studies Certificate as it relates to naval integration, a priority of senior naval leadership. Additionally, USNCC assessed a command nomination process to determine its effect on student participation, persistency, and achievement of learning outcomes. The command nomination process is an important development for the USNCC because it allows commands

Table 1. English course versus no course, performance comparison

	No English screening	English screening
Total students	185	1320
Below C- total	20	45
Total % below C-	10.8%	3.4%

Source: USNCC enrollment data.

to determine which of their sailors or Marines can become students to pursue advanced education based on the command’s evaluation of talent and the skills required for the unit. In addition to active-duty servicemembers, a portion of the students participating in the 3A pilot are reservists; their participation will allow USNCC to gain feedback from the reserve community, which up until the start of the 3A pilot was not possible.

The 3A pilot started on 28 May 2023 and concluded on 12 May 2024. This pilot allows students to start NAV101 and complete their Naval Studies Certificate through NAV105 within one year. Based on USNCC’s faculty and adjunct capacity, the college planned the 3A pilot for a maximum cohort of 250 sailors and Marines (to include 25 Navy and 25 Marine Corps reservists), with specific operational experience in a naval integrated environment, such as Expeditionary Strike Group, a Marine Expeditionary Force, an Amphibious Readiness Group, a Marine Expeditionary Unit, or similar task organizations.

Likely due to enlisted servicemember unfamiliarity with USNCC and the initial attempt at a command nomination process, only 61 students enrolled in NAV101 in May 2023 at the start of the 3A pilot. Even though this is fewer students than USNCC anticipated, valuable insight from students and commanders will help determine if the Naval Studies Certificate is relevant to the Navy and Marine Corps enlisted servicemembers. The 3A pilot introduced technological initiatives, such as the command nomination process to allow greater Fleet leadership interaction with the USNCC and a new command feedback mechanism to gain new insight on USNCC students and programs that may have greater application in future USNCC certificates and degrees.

Simulation and Wargaming in Naval Studies Courses

Simulation and wargaming (S&W) has been an influential aspect of the naval Services for more than 100 years. Currently, there is an S&W component at the Naval Academy, the Naval War College, the Naval Postgraduate School, and Marine Corps University. While no similar S&W program exists at the enlisted level, there is an opportunity for USNCC to be the leader in enlisted computer-driven S&W by developing and validating use of S&W in USNCC courses. USNCC has explored options to implement S&W in both NAV101 and NAV105 and has found that these courses could implement S&W rather well, especially multi-module assignments over a number of weeks.

Several factors support the inclusion of S&W within the naval studies program, including USNCC's institutional outcomes two through five and the guiding principle to be innovative.¹⁶ S&W within naval studies would also enhance USNCC's strategic goals, specifically Strategic Goal One: Develop and Implement Naval Relevant Programs and Outcomes (Objectives 1.1, 1.2, and 1.5) and Strategic Goal Five: Drive Impact for the Naval Services and Naval University System (Objectives 1.1, 1.2, and 1.5).¹⁷ S&W is also in line with the goals of the *Naval Education Strategy 2023*, specifically Ends: Education for Force Readiness and Competitive Advantage; LOE-3: Strengthen the NUS; and Objective 3.7: Implement Wargaming Best Practices.¹⁸

After preliminary conversations with an industry leading contractor, it is possible to use existing simulations for NAV101 and NAV105. For NAV101, two to three hours of added coursework is required, but it is manageable due to the current course organization and content. NAV101 would be supported by the contractor's international organizations simulation, which directly supports course learning objectives two, three, and four. The contractor has multiple South China Sea simulations that fit within the existing framework of NAV105.

¹⁶ "About USNCC," USNCC.edu, accessed 26 September 2024.

¹⁷ *United States Naval Community College Strategic Plan, 2022–2026* (Quantico, VA: USNCC, 2021), 11–15, 24–25.

¹⁸ *Naval Education Strategy 2023* (Washington, DC: Department of the Navy, 2023), 7, 14–15.

The addition of S&W would require the USNCC to develop or amend current assessment measures. A key metric that would be worth capturing is the impact that S&W has on student engagement. Based on the research conducted by scholars addressing the benefits of military S&W in an education environment, S&W can provide meaningful student engagement and expose students to different perspectives.¹⁹ This can be measured in assignment rubrics or through specifically designed survey questions. The USNCC has a chance to demonstrate adaptability and deliver unique education opportunities for the enlisted force. There is also risk in terms of budget and adopting unproven content into existing courses. This can be mitigated by adopting a pilot course to work out any difficulties on a small scale, evaluate the outcomes, ensure that it is offline compatible and see if the results are deemed acceptable, and implement for wider adoption.

As the USNCC continues to explore S&W implementation and FedRAMP requirements relevant to S&W adoption, the Krulak Center for Innovation & Future Warfare at Marine Corps University has extended an invitation for all USNCC students to participate in its multiple war-games.²⁰ When S&W is integrated into USNCC courses, USNCC faculty are keen to capture any lessons learned and student feedback through these opportunities, especially as they relate to exposing and familiarizing enlisted servicemembers with military wargames.

¹⁹ Peter P. Perla and Ed McGrady, “Why Wargaming Works,” *Naval War College Review* 64, no. 3 (2011): 111–30; P. C. Combe II, “Educational Wargaming: Design and Implementation into Professional Military Education,” *Journal of Advanced Military Studies* 12, no. 2 (2021): 115–38, <https://doi.org/10.21140/mcu.20211202004>; E. M. Walters, “Wargaming in Professional Military Education Challenges and Solutions,” *Journal of Advanced Military Studies* 12, no. 2 (2021): 81–114, <https://doi.org/10.21140/mcu.20211202003>; and A. M. Rosen and L. Kerr, “Wargaming for Learning: How Educational Gaming Supports Student Learning and Perspectives,” *Journal of Political Science Education* 20, no. 2 (2024): 318–35, <https://doi.org/10.1080/15512169.2024.2304769>.

²⁰ FedRAMP refers to the Federal Risk and Authorization Management Program, which is a government program that standardizes security assessment, authorization, and monitoring of cloud products and services.

Enhance USCG Components of Naval Studies Courses

Survey results from early naval studies courses suggested that the courses did not contain sufficient material on the role of the Coast Guard as a naval Service. In fact, Coast Guard participation is a fundamental element of the tri-Service maritime strategy, *Advantage at Sea: Prevailing with Integrated All-Domain Naval Power*.²¹ Few USNCC students from the Navy or the Marine Corps had worked with Coast Guard personnel and lacked an understanding of their duties and roles. Led by a retired Coast Guard captain, the faculty reviewed all five courses to assess where relevant Coast Guard material should be added to the courses. In four of the courses, faculty were able to add significant Coast Guard material to improve course coverage of all three Services. In NAV 101, faculty added ethical examples relevant to Coast Guard operations. In NAV 102, faculty added a section on the Coast Guard's operations during the Vietnam War. Faculty expanded the discussion in NAV 103 related to inter-Service relationships to include how the Coast Guard integrates into Joint operations. In NAV 105, faculty rewrote discussion board questions and added possible paper topics that focus on the relationship between the three Services in the international context. As reported in course-end surveys following the additions of new Coast Guard material, the new Coast Guard material improved the courses, led to stronger engagement by Coast Guard students, and helped Navy and Marine Corps students develop a better understanding of the roles and mission of the Coast Guard.

Alternate Delivery Tests

USNCC is committed to an asynchronous, online course delivery as its primary mode of teaching. But with the tempo and worldwide deployed student body that USNCC serves, the faculty have tested several other delivery modes as potential additions to the standard delivery mode to be able to reach students in circumstances where they cannot easily or reliably access online coursework. USNCC has tested a short course with partners from the Chief Digital and Artificial Intelligence Office (CDAO)

²¹ *Advantage at Sea: Prevailing with Integrated All-Domain Naval Power*, December 2020 (Washington, DC: Department of the Navy, 2020).

and the nonprofit educational organization Udacity. USNCC also conducted a test of low-bandwidth classwork for a ship underway. USNCC is also currently exploring other digital platforms that offer short courses and other nondegree credentials.

Chief Digital and Artificial Intelligence Office (CDAO) and Udacity

USNCC has experimented with alternative delivery methods to help determine student interest and possible areas where the USNCC could modify its offerings to reach more students. One such experiment was the offering of an artificial intelligence/machine learning (AI/ML) course in collaboration with CDAO and Udacity. USNCC handpicked 50 wait-listed applicants who all expressed interest in AI and had the required background in math. The selected applicants participated in an online CDAO course called Introduction to Machine Learning with TensorFlow, which started in May 2023 and ended in September 2023.²² This was a pilot offering designed to evaluate student interest, identify potential future course delivery methods, and explore collaboration with the CDAO. The course was completely online and offered asynchronously through Udacity's learning platform, allowing sailors, Marines, and coast guardsmen the opportunity to complete coursework during their off-duty time from anywhere provided they have a stable internet connection. However, by the end of September, only 1 of the 50 students completed the course.

USS Paul Hamilton

Another experiment intended to offer educational opportunities to deployed students was the *USS Paul Hamilton* (DDG 60) pilot course offering NAV101: Naval Ethics and Leadership. The goal of this experiment was to evaluate a low-bandwidth version of the course that did not require constant connectivity and peer interactions. The premise of this pilot course was to create an accessible version of NAV101 that even underway students could complete in a deployed operational tempo. The *Aegis*-class

²² The course is currently offered at "Introduction to Machine Learning with TensorFlow," Udacity.com, accessed 26 September 2024.

destroyer *Paul Hamilton* was selected to test the low bandwidth version of this course and modified procedures because this class of ship offers the best connectivity used by the smallest number of crew.

The *Paul Hamilton*'s education service officer (ESO) identified 17 students to participate, and while students had some difficulty downloading the three blocks of course material (weeks 1–3, 4–6, 7–8), they were able to ultimately complete the downloads successfully. Additionally, students were given a 4-week goal to complete each section for a total of 12 weeks, adding additional time to the standard 8-week version of the course as offered online to all other students. The course instructor encouraged students to work at their own pace or request extensions if the operational tempo became too demanding. The only time students were required to log into the learning management system was to submit their work, but they also had the option of turning in work via email. Despite multiple communication attempts by the faculty through the management system, only one student submitted one assignment between May 2023 and October 2023. Even when the *Paul Hamilton* returned to home port, students did not turn in another assignment.

Based on limited student feedback from the one student who completed the course, a plausible reason for the poor completion and participation rates for both the CDAO/Udacity and *Paul Hamilton* courses is an underestimation of the importance of a “course community” and the camaraderie created by sharing ideas. Regardless of the course material and delivery method, an engaged community of learners may be required to maximize course participation and engage in critical thinking. USNCC plans to conduct another underway experiment in 2024 with a different ship and will test improvements to community-building (including predeployment course training for student mentors and/or the ship's ESO), more structured assignments, and a redesign of the course modules. By creating structured instructor-student touchpoints in each of the three course blocks and designing student interaction assignments in conjunctions with the course's short answer assignments, students will have a more opportunities to engage with their instructor and fellow students.

Pre- and Nondegree Credentials

USNCC provides supplementary predegree and nondegree educational opportunities through coordination with academic partner institutions or industry online platforms to offer, when available, relevant credentials, professional certificates, licenses, and qualifications within the curriculum matrix. These additional curricular offerings must add skills and knowledge beyond what enlisted members receive through general military training (GMT) badges that servicemembers earn during their annual mandatory online training sessions. Servicemembers are known to sometimes “click through” these GMT trainings as quickly as possible, therefore lessening the educational and even sometimes the training value of these modules. Additionally, any predegree and nondegree credentials, professional certificates, licenses, and qualifications must be “naval relevant” and useful for both the servicemember and the servicemember’s command or organization. Examples of additional curricular offerings that meet these criteria might be CompTIA (for IT careers) certificates, a Remote Pilot Certificate from the Federal Aviation Administration (FAA), preparation for FAA Aviation Maintenance Technician Certifications, and the previously discussed Naval Studies Certificate.

Scholarly research indicates that these predegree and nondegree credentials should improve student engagement and student completion rates as they can offer tangible academic and professional milestones earlier in the servicemember’s pathway toward their associate of science (AS) or associate of applied science (AAS) degrees. The research also suggests that predegree and nondegree offerings should increase command support for these academic pathways through organic development of a more competent and qualified workforce available to the sponsoring military unit.²³ In exploring delivery mechanisms and modalities for these pre-

²³ Sean R. Gallagher, *Educational Credentials Come of Age: A Survey on the Use and Value of Educational Credentials in Hiring* (Evanston, IL: Center for the Future of Higher Education and Talent Strategy, Northeastern University, 2018); John Brennan et al., *Study on Innovation in Higher Education: Final Report* (Luxembourg: Publications Office of the European Union, European Commission Directorate for Education and Training Study on Innovation in Higher Education, 2014); and David P. Baker, “Forward and Backward, Horizontal and Vertical: Transformation of Occupational Credentialing in the Schooled Society,” *Research in Social Stratification and Mobility* 29, no. 1 (2011): 5–29, <https://doi.org/10.1016/j.rssm.2011.01.001>.

degree and nondegree credentials, the U.S. Air Force's Digital University served as one currently functioning military model of online curriculum delivery. The platform demonstrates capacity for delivery of online educational coursework.²⁴

USNCC is currently exploring the technical and economic efficiencies of joining with an existing learning management system (LMS) as opposed to creating a unique delivery platform built “ground up” for USNCC use. The four primary concerns applicable to USNCC use cases include: 1) access to the LMS from deployed locations around the globe; 2) student identification verification in an asynchronous, online environment; 3) measurable testing and assessment to ensure mastery of the material; and 4) platform cost per student credit hour.

Preliminary Findings

Results (where available) for each experiment offered some insight into the utility of each of the experiments; broader results about student performance and satisfaction with their academic programs also suggest that USNCC's academic efforts are helping the college achieve its mission to support the Services.

Two key questions for the faculty as it established USNCC's academic programs are: What do the students know, and how do they know it? Some simple indicators of success offer rough answers to these questions, such as responses on student end-of-course surveys, substantial reenrollment or student persistence, or increasing enrollment numbers every term.²⁵ Student responses have been favorable for USNCC programs so far, indicating satisfaction with the relevance of the programs:

Students perceived that the programs are delivering valuable, applicable skills that align with their lifelong learning aspirations. Of the 865 respondents, 72% of students perceive that learned skills in USNCC courses are Relevant or Highly Relevant to their future college plans.

²⁴ “Digital University,” U.S. Air Force, accessed 1 March 2024.

²⁵ Noelia Pacheco-Díaz et al., “Faculty Use of End-of-Course Evaluations,” *International Journal of Teaching and Learning in Higher Education* 33, no. 3 (2022): 285–97.

Table 2. USNCC student persistence rates, by term

Pilot II start	Summer 2022	Fall I 2022	Fall II 2022	Spring I 2023	Spring II 2023	Summer 2023
Term 1	100%	100%	100%	100%	100%	100%
Term 2	96%	90%	87%	75%	58%	73%
Term 3	93%	80%	63%	74%	52%	
Term 4	90%	66%	43%	64%		
Term 5	78%	49%	39%			
Term 6	64%	39%				
Term 7	52%					

Source: *State of the College* 2023, 8.

While 82% believe that learned skills in USNCC course are Relevant or Highly Relevant to their role as enlisted service members. More broadly, 84% Agree or Strongly Agree that their programs and courses as naval-relevant. And more granularly, 90% (n = 588) of students Agree or Strongly Agree that they are better able to address ethical issues and consequences.²⁶

As a rough measure of USNCC’s performance against national averages, “Only 26% of working U.S. adults with college experience strongly agree that their education is relevant to their work and day-to-day life,” demonstrating that USNCC’s courses matter to their students.²⁷

Student persistence has also been strong, coming in higher than the national average of 43.8 percent for part-time students enrolled in two-year institutions.²⁸ Persistence rates for USNCC students are reported in table 2.

Student performance in their programs, and in the Naval Studies Certificate, is another important datum that the USNCC faculty track.

²⁶ *State of the College* 2023, 10.

²⁷ *From College to Life: Relevance and the Value of Higher Education* (Washington, DC: Strada Education Fund and Gallup, 2018), 3.

²⁸ *State of the College* 2023, 8.

USNCC partner schools submit an annual assessment report covering the performance of all students, including USNCC students, in the programs they operate in consortium with USNCC. USNCC's institutional effectiveness office surveys and assesses all Naval Studies courses, as well. Faculty use the same rubrics for assessing course and program learning outcomes in all Naval Studies courses, allowing them to track development of student skills throughout the program.

An example of this approach tracks student performance in achieving USNCC's third program learning outcome (PLO), which states:

Students shall demonstrate 21st Century skills inclusive of critical thinking, analytical reasoning, information literacy, data fluency, and communicative fluency (inclusive of writing and verbal communications).²⁹

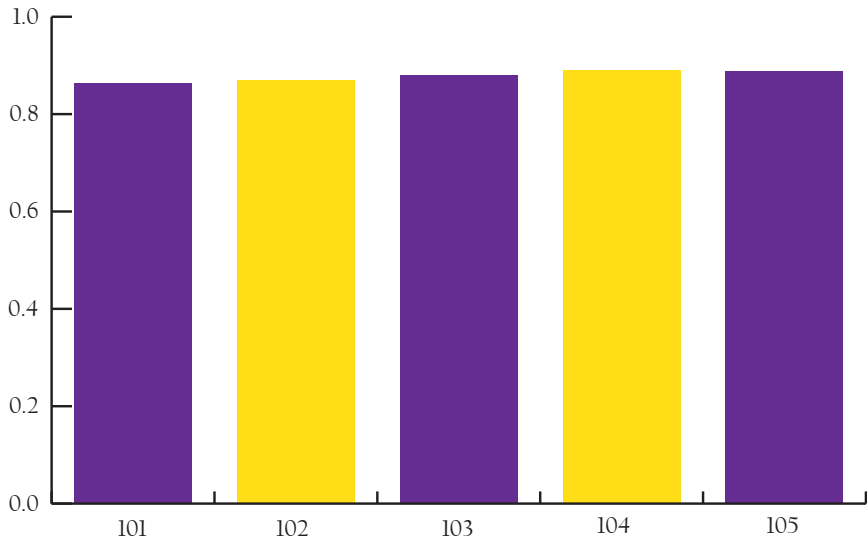
Figures 1 and 2 report student assessment outcomes for this PLO. The trend lines indicate that student performance on the PLO does improve over time, as expected. Forum scores in NAV 101 weigh more heavily than in later naval studies courses because there is more emphasis on written work in NAV 102–105. The data shows that student skills in communications, critical thinking, and analytical reasoning all grow over the course of the Naval Studies Certificate program.

Future Plans: AI in the Classroom

In the near-term future, USNCC plans to implement the use of AI in curricular design, coursework design, and classroom teaching environments. Soon-to-be-released software programs and updates such as Anthology's Blackboard Learn and Coursera AI functionality, among others, will facilitate improved academic program design, course content and learning objective assessment development. The use of generative AI (GenAI) within asynchronous instructional modules should soon allow for real-time lesson adaptation and coaching based on student competencies and performance. Use cases where this AI-driven adaptability could be particularly groundbreaking are mathematics and computer language programming

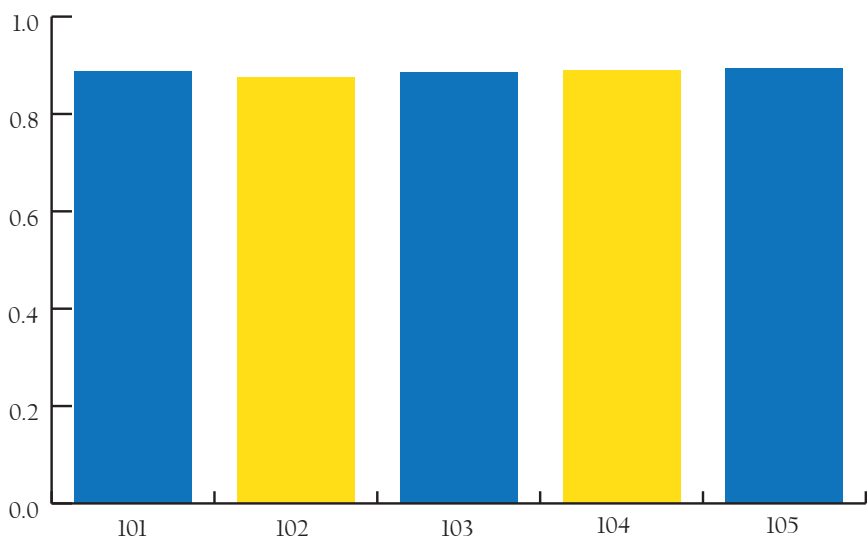
²⁹ 2023–2024 USNCC Academic Catalog (Quantico, VA: USNCC, 2023), 10.

Figure 1. Average of PO3 rubric assessment, by course (written, bibliography, oral presentation) rubrics



Source: USNCC Office of Institutional Research, Assessment, and Effectiveness.

Figure 2. Average of PO3 rubric assessment, by course (including forum rubrics)



Source: USNCC Office of Institutional Research, Assessment, and Effectiveness.

tutoring or eventually adaptive English and/or writing tutoring. Ultimately, many predict that GenAI might allow for near-real-time essay grading and assessment.

In terms of improved naval relevance, USNCC maintains regular and meaningful collaboration with responsive Navy, Marine Corps, and Coast Guard fleet subject matter experts and enlisted leaders embedded within training and operational forces. USNCC has established advisory groups for curricular relevance for specific program areas such as leadership coursework, STEM (science, technology, engineering, and mathematics) coursework, operational support coursework, and emerging technology coursework to enable easier communications between the operational forces and USNCC. USNCC faculty and staff, particularly the associate deans and the command senior enlisted leader conduct site visits to military units around the globe to ensure that USNCC programming remains closely aligned with evolving fleet operations. Additionally, USNCC faculty and staff regularly review military training programs to evaluate their ability to award transfer credits to academic transcripts while further aligning USNCC courses and academic programs to Service needs.

Future Plans: New Program Launches

As USNCC looks to the next two years, it plans to launch new academic degree programs in several fields identified by DON as vital to help the Services master future technological advancements that will be (or are already) deployed to units, ships, and aircraft. In 2024, USNCC plans to launch competency-based education (CBE) programs in select engineering and health sciences fields to offer multiple delivery options to its students. USNCC's intent is to mirror its asynchronous online programs with CBE versions for students who are attuned to the self-directed, speedier approach offered in the CBE model. In addition, DON has identified engineering programs as key for the future of USNCC; the college is soliciting partners for a range of engineering degrees (both AS and AAS) that would meet specific community needs across the naval Services. Ultimately, USNCC expects to offer programs in electrical engineering, communications/electronic warfare systems, software engineering/technology, industrial engineering, industrial maintenance technology, control (sys-

tems) engineering/technology, and marine (ocean) engineering/marine technology. The college also plans to launch a program in emergency management, which is of particular interest to the U.S. Coast Guard, given its missions.

Conclusions

Experiments at USNCC have taught the staff and faculty many important lessons and point the way to continuing the effort at innovation to keep courses interesting and to add new programs and delivery modes as appropriate to maintain the relevance of USNCC programs to students and to the fleet. Constant experimentation *in courses* helps sustain engagement of both the students and faculty. As the research indicated, a sense of community is a key ingredient to student success, and it is harder to build and sustain that in online programs than in face-to-face environments.³⁰

Alternate *delivery modes* take practice, planning, and more time than expected. USNCC faculty realized while testing each of the alternate ideas that there were many more moving parts than initially anticipated, and that setting up and executing a new delivery mode was far more time-consuming than estimated. Future tests will benefit from realizing that a longer, more comprehensive planning phase must precede any new delivery mode launch.

Connection to the force helps with relevance and must be maintained and strengthened. USNCC's mission is directly focused on supporting the Services with relevant programs and courses, and the best way to ensure ongoing relevance is to report on USNCC programs to advisory groups and senior enlisted leaders in the force and solicit their advice and recommendations to stay on track.

The USNCC program model grants its students access to established, expert programs, but comes with administrative complexity. Every launch takes longer than expected because every partner that joins

³⁰ Archambault, Leary, and Rice, "Pillars of Online Pedagogy"; Lewis, "Best Practices for Improving the Quality of the Online Course Design and Learners Experience"; Shin and Cheon, "Assuring Student Satisfaction of Online Education"; and Steele, Holbeck, and Mandernach, "Defining Effective Online Pedagogy."

the USNCC consortium has its own internal processes and accreditation requirements that vary from institution to institution and sometimes within the colleges of a larger partner. USNCC faculty and academic staff must coordinate every program with partners, as well as other offices at USNCC. Much like launching alternate delivery tests, every program launch is a complicated, time-consuming process that has to be done correctly to ensure that the best programs are available to USNCC's students.

Enlisted members are hungry for higher education, and USNCC has to meet that hunger with relevant, career-enhancing options that also improve the naval Services' warfighting abilities. USNCC's dedication to experimentation and data-driven decision-making is all about meeting this mission head on.

USNCC expects to launch additional programs in 2025, including emerging technologies, health sciences, AI/ML, robotics and mechatronics, and additive manufacturing. As USNCC expands its programs to serve its target student body of up to 25,000 enrolled students, it will continue to experiment with credit and noncredit programs, with improvements in the naval studies courses, and with alternative delivery modes, all in the pursuit of serving its students as well as possible, and to achieve its mission to enhance readiness, improve warfighting ability, and set up its students for successful lifelong learning.



CHAPTER 11

Oral Comprehensive Exams

Lessons on the Assessment of Student Achievement
of Program Learning Outcomes at the U.S. Army War College

*Colonel Christopher J. Hickey (USA), PhD;
Professor Brett D. Weigle, EdD; and
Professor Matthew C. Woessner, PhD*

This chapter provides helpful information on the development and execution of comprehensive examinations in the context of professional military education (PME). However, it is also relevant to civilian institutions. This is important as programs shift to an outcomes-based model requiring student and program assessment forms that differ from traditional PME approaches, such as calculating time spent on specified topics. The observations made in this chapter are relevant to any educational setting that seeks to develop professional judgment regardless of the students' civil or military status or their military rank and Service.

The chapter begins by explaining why leaders and faculty in PME may want to implement these examinations. It reviews the literature on comprehensive examinations in undergraduate and graduate education. Then, it offers a short case describing the evolution of comprehensive exams to assess student learning at the U.S. Army War College (USAWC) during a decade. In keeping with the U.S. Department of Defense's (DOD) increasing emphasis on data-informed decision making, the chapter describes and evaluates data analysis and linear regressions designed to test the effectiveness of the comprehensive exam in two areas: measuring student performance and the effectiveness of specialized academic programs in increasing student performance. The chapter ends with observations

with clear implications for military and civilian graduate education in various educational settings.

Literature on Comprehensive Exams

There is surprisingly robust and longstanding literature on comprehensive examinations in graduate-level education. One can argue that it has its roots in the tutorial model of education used at Cambridge and Oxford Universities in England and carried on in early American universities such as Yale and Harvard. Small groups of students studied with a faculty member, preparing for frequent discussion sessions that should be “demanding, stimulating and thought-provoking, for student and tutor alike.”¹ In 1931, William H. Spencer described the implementation of comprehensive examinations at the University of Chicago’s School of Commerce and Administration in ways that remain relevant today. The school’s intention to implement the examination was broadly similar to today’s concept of outcome-based educational assessment, with the award of a degree being “based upon attainments as tested by comprehensive examinations rather than upon course requirements and course examinations.”² Among the reasons Spencer offered for requiring a series of comprehensive examinations was developing in the student a wider understanding of the field of knowledge than would come from individual courses and providing the student and school with a better understanding of the student.³ He noted that the school’s earlier attempt to implement comprehensive exams had failed because it had difficulty developing appropriate questions and evaluation criteria.⁴

Rodney T. Hartnett and Warren W. Willingham showed in 1980 that establishing an appropriate process and criteria for measuring adult educational success has long been a critical challenge for educators. Their research suggested that many schools had yet to articulate to students and

¹ David Mills and Patrick Alexander, “Are Oxbridge Tutorials Still the Best Way to Teach Students How to Think?,” *Conversation*, 8 July 2015.

² William H. Spencer, “Comprehensive Examinations in Business,” *Journal of Business of the University of Chicago* 4, no. 3 (1931): 98.

³ Spencer, “Comprehensive Examinations in Business,” 98, 100.

⁴ Spencer, “Comprehensive Examinations in Business,” 100.

faculty the purpose, process, preparation, and evaluation mechanisms of their comprehensive examinations despite decades of literature describing this problem. They warned that “the suggestion that graduate faculties should specify the competencies they expect of students and construct examinations to test whether those competencies have been achieved is seldom given serious consideration, in spite of the success of such practices in several professional fields.”⁵ Tellingly, although they stated that the paper “attempted a general analysis of the strengths and weaknesses” of several student performance evaluation methods, they had essentially nothing positive to say about how comprehensive examinations actually were being conducted.⁶

In fact, much of the literature deals with the less positive aspects of comprehensive examinations from the perspective of students, faculty, and institutions. Wayne P. Anderson and his colleagues argued in 1984 “that if students can see a clear reason for the examination and have a model to follow in preparation, they will learn more, and their test anxiety will be reduced.”⁷ Less clearly stated is the implication that, too often, this was not occurring, leading to missed learning opportunities and frustrated students. Their review of the literature on comprehensive exams suggested that the three main reasons schools conducted comprehensive exams were to screen out underperformers, impose a rite of passage, and provide a chance for students to organize and integrate their knowledge. They proposed that the third reason is the one that leads to the best experience. They described a deliberate system used in their school to clarify that the event is about integrating knowledge and helping students prepare. The faculty coached the students as they wrote answers to previously used questions about six weeks prior to actually taking the exam.⁸ This made it clear what was expected and made the

⁵ Rodney T. Hartnett and Warren W. Willingham, “The Criterion Problem: What Measure of Success in Graduate Education?,” *Applied Psychological Measurement* 4, no. 3 (1980): 284–85, <https://doi.org/10.1177/014662168000400301>.

⁶ Hartnett and Willingham, “The Criterion Problem,” 284–85, 289.

⁷ Wayne P. Anderson et al., “Reasons for Comprehensive Examinations: A Re-evaluation,” *Teaching of Psychology* 11, no. 2 (1984): 78, https://doi.org/10.1207/s15328023top1102_4.

⁸ Anderson et al., “Reasons for Comprehensive Examinations,” 78–80.

actual exam the direct culmination of the preparation process with little wasted effort.

Teri Olisky Loughead argued in 1997 that Benjamin S. Bloom's *Taxonomy of Educational Objectives* provides a valuable tool in the "process of designing, preparing for, evaluating, and providing feedback."⁹ Her research showed significant problems with all of these areas across academia.¹⁰ She demonstrated that Bloom's taxonomy offered a simple common logic that sought "increased coherence within the profession concerning the purposes and processes" for comprehensive examinations.¹¹

Writing in 2014, Susan K. Burke and John T. Snead echoed Wayne P. Anderson et al. and Loughead in demanding that useful exams "require students to integrate, analyze, and synthesize knowledge."¹² They specifically warned that "coherent educational objectives, as opposed to ritual and tradition," are best practices for comprehensive exams.¹³

Scholars continue to find evidence that comprehensive examinations are being conducted in ways that do not achieve their assessment intent and may cause students unnecessary stress. For example, in 2020, Gillian Stewart-Wells and K. Mallery Keenan found that "there is ample discussion in the literature that comprehensive exams are frustrating and fruitless to both students and faculty" in doctoral programs.¹⁴ In 2023, Australian researchers Sarah French, Ashton Dickerson, and Raoul A. Mulder reviewed 122 academic studies concerning "high-stakes examinations . . . as summative assessments in higher education" for undergradu-

⁹ Benjamin S. Bloom, *Taxonomy of Educational Objectives* (Boston, MA: Allyn and Bacon, 1956); and Teri Olisky Loughead, "The Doctoral Comprehensive Examination: Fine-Tuning the Process," *Counselor Education and Supervision* 37, no. 2 (1997): 140, <https://doi.org/10.1002/j.1556-6978.1997.tb00539.x>.

¹⁰ Loughead, "The Doctoral Comprehensive Examination," 140–47.

¹¹ Loughead, "The Doctoral Comprehensive Examination," 147.

¹² Susan K. Burke and John T. Snead, "Faculty Opinions on the Use of Master's Degree End of Program Assessments," *Journal of Education for Library and Information Science* 55, no. 1 (2014): 28.

¹³ Burke and Snead, "Faculty Opinions on the Use of Master's Degree End of Program Assessments," 28.

¹⁴ A. Gillian Stewart-Wells and K. Mallery Keenan, "Assessing Doctoral Students: A Background on Comprehensive and Authentic Assessments," *Journal of Continuing Higher Education* 68, no. 2 (2020): 85, <https://doi.org/10.1080/07377363.2020.1743950>.

ates.¹⁵ They found that support for their use “is largely rooted in opinion and pragmatism” with little evidence of “pedagogical merit.”¹⁶

Both groups of researchers favor an approach based on “authentic assessment,” which seeks to “evaluate individuals in more real life or genuine contexts.”¹⁷ They propose a system of “benchmarks” where each task is closely related to the student’s future career. The student completes the tasks, such as creating curriculum material or scholarly research, to demonstrate the integration of the program learning outcomes.¹⁸ This approach has also been titled *ungrading*, where formative assessments build knowledge throughout a course with no numerical marks or grades but adjusted via extensive faculty feedback. The student aggregates their work into a comprehensive portfolio that receives a summative graded assessment for a holistic look at the student’s development rather than accumulated interim grades from quizzes, homework, and exams.¹⁹

A positive contemporary example of summative comprehensive examinations comes from the U.S. Army’s Command and General Staff School (CGSS). Researchers describe the CGSS case study to show the application of best practices in developing and conducting comprehensive exams. Employing Army and Joint Staff PME guidance, the faculty constructs a set of learning outcomes nested at the institutional, course, and lesson levels, then “creates and presents the curriculum organized to meet these outcomes and plays the central role in the creation of appropriate assessment instruments to determine student performance and provide

¹⁵ Sarah French, Ashton Dickerson, and Raoul A. Mulder, “A Review of the Benefits and Drawbacks of High-Stakes Final Examinations in Higher Education,” *Higher Education* 88 (December 2023): 1, <https://doi.org/10.1007/s10734-023-01148-z>.

¹⁶ French, Dickerson, and Mulder, “A Review of the Benefits and Drawbacks of High-Stakes Final Examinations in Higher Education,” 1.

¹⁷ Stewart-Wells and Keenan, “Assessing Doctoral Students,” 85.

¹⁸ Stewart-Wells and Keenan, “Assessing Doctoral Students,” 87–98.

¹⁹ See, for example, Horace T. Crogman et al., “Ungrading: The Case for Abandoning Institutionalized Assessment Protocols and Improving Pedagogical Strategies,” *Education Sciences* 13, no. 11 (2023): 1091, <https://doi.org/10.3390/educsci13111091>; and Robert Talbert, “What I’ve Learned From Ungrading,” *Inside Higher Ed*, 26 April 2022.

data for curricular improvement.”²⁰ Michelle A. Miller, James Pennington, and Billy D. Miller Jr. describe it:

The online exam is an objective test administered via Blackboard [learning management system], consisting of fifty knowledge and comprehension-level questions. These questions are prerequisites of what every graduate needs to know. The oral board is a grueling one-hour individual session conducted by a panel of three skilled faculty members. The oral faculty board uses four integrated questions to assess the student’s mastery of the core material. Blackboard allowed as many as 1,200 users to take the written exam simultaneously and provided a means for the faculty to manage the assessment of so many exams in a reasonable time.²¹

The comprehensive exam is explained to the CGSS students at the beginning of the course, and they are provided with a set of sample questions, some of which will be on the exam. Echoing the findings of Anderson et al., preparation for the examination becomes a routine and deliberate part of the program as students are aware of the desired outcomes, the examination process, and the need to integrate material from various courses and experiences.

After the comprehensive examination assesses student knowledge and integration of concepts, they apply this material in the program’s second half, focused on military planning and operations. The school has found additional value in the examinations as a data source to improve the educational program and provide evidence of student achievement in joint learning areas during accreditation by the Joint Staff. A key fac-

²⁰ *Officer Professional Military Education Policy*, Chairman of the Joint Chiefs of Staff Instruction 1800.01F (Washington, DC: Joint Chiefs of Staff, 2020); and Ann Collins Royce, Ryan Welch, and James B. Martin, “Assessment and Evaluation Practices in Army Military Education,” in *Assessment, Evaluation, and Accountability in Adult Education*, ed. Lilian H. Hill (New York: Routledge, 2020), 75–76, <https://doi.org/10.4324/9781003443117>.

²¹ Michelle A. Miller, James Pennington, and Billy D. Miller Jr., “U.S. Army Increases Rigor at the Command and General Staff College,” *Command and General Staff College Foundation*, 8 February 2019.

tor the authors identify in the success of assessments at CGSS is the 4:1 student-to-faculty ratio and the existence of relevant institutional experts on the staff. The faculty and staff capacity and expertise make a rigorous, well-structured assessment process possible.²²

This chapter adds to the existing literature in three ways. First, it provides a study of the evolution of a successful comprehensive exam process at the USAWC that deliberately sought to overcome the challenges described in this literature. Second, the chapter uses linear regression to analyze anonymized data from three years of exams to address effectiveness issues. Third, it offers observations that have broad implications for educators in military and civilian graduate education in various settings.

Why Administer Comprehensive Exams in PME?

Given the pessimistic tone of the literature, one may wonder why PME leaders would consider comprehensive examinations when deciding how to invest the time and attention of their faculty, staff, and students. Multiple factors in the PME environment support the value of the examination process to the students and the justification of the institutional investment.

The most important reason is that PME strategic guidance and professional consensus at the USAWC lead to a focus on outcomes. Anna T. Waggener, a former director of institutional assessment at the USAWC, presents a helpful 2015 review of the congressional interest in Joint professional military education (JPME) dating back to 1975.²³ Among other points she makes in her overview is that policy decisions dating back decades set the high resource levels for faculty to student ratios and linked classroom activities, resourcing, and strategic guidance in ways that have created the complex but successful system of PME we know today. Strategic guidance documents have long emphasized the importance of high-quality assessment of student learning and program effectiveness.

²² Royce, Welch, and Martin, "Assessment and Evaluation Practices in Army Military Education," 75–79.

²³ Anna T. Waggener, "Joint Professional Military Education: A Retrospective of the Skeleton Panel," *Joint Force Quarterly* 77, no. 2 (2015): 55–59.

In 2020, the seven senior generals and admirals who lead the U.S. military Services published *Developing Today's Joint Officers for Tomorrow's Ways of War*. The first critical task it established was “Adapt and Innovate PME,” which included the direction to “shift our PME curricula from a predominately topic-based model to an outcomes-based approach and emphasize ingenuity, intellectual application, and military professionalism in the art and science of warfighting, while deepening knowledge of history.”²⁴

The 2020 version of the *Officer Professional Military Education Policy* (OPMEP) made significant changes from earlier versions to implement *Developing Today's Joint Officers for Tomorrow's Ways of War*. The most important change was the direction that “shifts officer JPME to an OBME [outcomes-based military education] approach with emphasis on direct assessments of student achievement of program learning outcomes.”²⁵ OBME is defined as “clearly focusing and organizing everything in an educational system around what is essential for all students to be able to do successfully at the end of their learning experiences. This means starting with a clear picture of what is important for students to be able to do, then organizing curriculum, instruction, and assessment to make sure this learning ultimately happens.”²⁶ The OPMEP builds on the scholarly concept of *authentic assessment* and directs PME institutions to build assessments that are a reasonable proxy for the challenges that will be faced by the student body, not just in their next job but in the rest of their professional life. This policy requires PME schools to ensure that “OBME assessments will approximate conditions under which the graduate would be expected to achieve the same outcomes in the operational environment (hereafter referred to as authentic assessments).”²⁷ Delivering well-reasoned answers to complex questions under pressure is an authentic expectation of USAWC graduates who might, for example, brief strategic leaders on a complex issue, testify before a congressional

²⁴ *Developing Today's Joint Officers for Tomorrow's Ways of War: The Joint Chiefs of Staff Vision and Guidance for Professional Military Education & Talent Management* (Washington, DC: Joint Chiefs of Staff, 2020), 5.

²⁵ *Officer Professional Military Education Policy*, 3.

²⁶ *Officer Professional Military Education Policy*, GL-6.

²⁷ *Officer Professional Military Education Policy*, A-5.

committee, or collaborate with senior officials in an interagency or multinational meeting.

Outcomes-Based Military Education Procedures for Officer Professional Military Education describes how “OBME shifts emphasis from focusing on compliance with mandating topics in the JPME curriculum to focusing on students achieving approved Program Learning Outcomes (PLOs). . . . JPME programs receive full certification under OBME based on evidence of compliance with Officer Professional Military Education Policy (OPMEP) standards for education and effectiveness of program learning.”²⁸ Therefore, it is not enough to measure what topics were taught; instead, the institution must find a way to assess students’ educational outcomes, as Kristin Mulready-Stone argued in “A New Form of Accountability in JPME: The Shift to Outcomes-Based Military Education.”²⁹

The new strategic guidance on PME is not restricted to commissioned officers. In 2021, the Senior Enlisted Advisor to the Chairman (SEAC) and the Service Senior Enlisted Advisors (SEAs) signed *Developing Enlisted Leaders for Tomorrow’s Wars: Our Shared Vision for Enlisted Professional Military Education & Talent Management*. This document mirrored the Joint Chiefs of Staff guidance on methods and expanded the degree of operational- and strategic-level knowledge and judgment expected of senior enlisted leaders.³⁰ They subsequently issued the *Enlisted Professional Military Education Policy* with implementation guidance tailored to the needs of future senior enlisted leaders. Everything in the guidance on officer education that leads to the need for authentic and cumulative assessments is mirrored in the guidance on senior enlisted leader education policy.³¹ This chapter may be of particular interest to leaders in senior enlisted education as they assist their students when transitioning into the many different and complex roles senior enlisted advisors serve in across the defense enterprise.

²⁸ *Outcomes-Based Military Education Procedures for Officer Professional Military Education*, CJCS Manual 1810.01 (Washington, DC: Joint Chiefs of Staff, 2022), 1.

²⁹ Kristin Mulready-Stone, “A New Form of Accountability in JPME: The Shift to Outcomes-Based Military Education,” *Joint Force Quarterly* 112, no. 1 (2024): 30–38.

³⁰ *Developing Enlisted Leaders for Tomorrow’s Wars: Our Shared Vision for Enlisted Professional Military Education & Talent Management* (Washington, DC: Joint Chiefs of Staff, 2021).

³¹ *Enlisted Professional Military Education Policy*, CJCS Instruction 1805.01C (Washington, DC: Chairman of the Joint Chiefs of Staff, 2021).

Within the U.S. Army, PME is viewed as an essential component to the development of the future leaders of the military profession. Therefore, CGSS and USAWC have been resourced on a scale that allows them to overcome many of the problems cited in the literature. As described in the case of CGSS, these graduate-level PME institutions enjoy four advantages over many of their civilian counterparts.³² They generally have a low student-to-faculty ratio, excellent physical facilities with extensive information technology support, a supporting staff familiar with detailed planning and assessment, and the ability to incentivize civilian and military faculty toward a relatively higher commitment of effort toward teaching and institutional service versus publishing than is the norm in civilian higher education.³³ The extent to which this is true at other PME institutions will influence their decision making on whether and how to implement comprehensive examinations or comparable assessments. Conversely, if resources are insufficient at a PME institution to support such a process, this may be one indicator that the institution is underresourced relative to the level of ambition in the guidance documents on strategic education.

A well-designed and executed comprehensive examination offers students a chance to synthesize an answer, showing that they understand and can employ critical concepts and material across the curriculum. It also allows them to demonstrate that they have developed thinking skills appropriate for their role as senior leaders or advisors to senior leaders, with responsibilities that span the levels of war. The examination offers the students a forum to showcase their improved ability to think deeply about a complex issue and then communicate concisely, clearly, and compellingly while under pressure and constrained by limited time.

Case Study at the U.S. Army War College

This section has four parts. First, it will define the program learning out-

³² Miller, Pennington, and Miller Jr., “U.S. Army Increases Rigor at the Command and General Staff College”; and Royce, Welch, and Martin, “Assessment and Evaluation Practices in Army Military Education.”

³³ *AY24 Resident Education Program Student Catalog* (Carlisle Barracks, PA: U.S. Army War College, 2023).

comes (PLOs) that the comprehensive exams are intended to assess. Second, it will outline the learning environment in the resident education program at the USAWC. Third, it will describe the leaders' intent for the exams. Finally, it will discuss the preparation for, conduct of, and assessment of the exams.

Program Learning Outcomes

Since 2016, the USAWC School of Strategic Landpower has constructed its curriculum around six program learning outcomes to produce graduates who exhibit desired leader attributes established in the OPMEP. The PLOs have been refined periodically but remain broadly consistent:

1. Discern the military dimension of a challenge impacting national interests; assess and communicate options for employing landpower in Joint warfighting and integrating with other instruments of national power.
2. Apply analytical frameworks and theories to analyze and evaluate strategies to address current and anticipated national security challenges across the competition continuum in a global environment.
3. Analyze military requirements of current and future security environments and how senior leaders use DOD processes and organizations to develop, integrate, and sustain necessary military capabilities.
4. Apply strategic leadership to address problems of organizational performance, adaptation, integration, and learning.
5. Lead teams with expert knowledge, develop professional networks, and collaborate with others to provide innovative solutions to complex, unstructured problems.
6. Apply strategic thinking and communication to analyze and evaluate the scope and implications of current and future national security and strategic military challenges.³⁴

³⁴ AY24 Resident Education Program Student Catalog, 6.

For the scope of the regression analysis (2021–23), the comprehensive exam assesses student attainment of PLOs 1, 2, 3, 4, and 6. The fifth PLO addresses interpersonal skills, networking, and collaboration at the senior level. The faculty committee responsible for the comprehensive exams does not believe the exam format is suited to assess this objective. Faculty use other tools to assess this outcome during seminar discussions and exercises.³⁵

The Army War College Resident Education Program

The students are organized into 24 seminars, with about 16 students from all Services of the U.S. armed forces (Active, Reserve, and National Guard) and international officers from more than 70 partner nations. Several U.S. government civilian agencies also send students to the program. Each seminar is comprised mainly of Army students, along with members of the aerospace and maritime Services and civilian students.³⁶

Each seminar's education is facilitated by a dedicated faculty team of three civilian and military professionals representing the three teaching departments in the resident program. The faculty teach a foundations course plus seven or eight core courses: 1) Strategic Leadership (SL); 2) Theory of War and Strategy (TWS); 3) National Security Policy and Strategy (NSPS); 4) Military Strategy and Campaigning (MSC); 5) Regional Studies Program (RSP); 6) Defense Management (DM); 7) China Integrated Course; and, for international students, 8) the Unified Command Course. There are additional requirements not discussed here, including a strategic research requirement, four electives, lectures, conferences, staff rides, speaking engagements, and Service fitness requirements. Beginning with academic year 2027, a new assessments-based tailored education (ABTE) model will further refine these requirements to the individual student.

Additionally, students may volunteer to apply for selection to participate in one of three competitively selected “immersive specialization” seminars instead of the seminar approach described above. The School

³⁵ AY24 *Resident Education Program Student Catalog*, 7.

³⁶ AY24 *Resident Education Program Student Catalog*, 5.

of Strategic Landpower dean describes them as “tailored versions of the core curriculum together with specially designed material, to achieve the PLO through the lens of a specific element of the strategic-enterprise level.”³⁷

- Advanced Strategic Art Program (ASAP)
- Carlisle Scholars Program (CSP)
- National Security Policy Program (NSPP)

These students take the same comprehensive exam as those in the non-specialized seminars.

Dean’s Intent for Comp Exams

The USAWC *Resident Program Academic Year 2024 Student Catalog* establishes policy for the current comprehensive exam:

Each student will participate in a comprehensive oral examination at the end of the core curriculum as one measure to assess his or her attainment of the Program Learning Outcomes. The exam will be conducted by members of a core faculty instructor team other than the team that taught that student’s seminar. The exam will consist of a dialogue during which the student will display his or her ability to:

- synthesize an answer that demonstrates command of key concepts and integrates material across the curriculum;
- demonstrate thinking skills with a strategic perspective (creative, systems, ethical, historical thinking), while considering other viewpoints (critical thinking);
- organize thoughts and articulate complex ideas with clarity and conciseness using appropriate frames of reference.

Every student must achieve a grade of B or higher

³⁷ AY24 *Resident Education Program Student Catalog*, 40.

to successfully complete the exam. Those students who receive a B– or lower on the exam will be offered a second opportunity. Those who do not succeed a second time will be referred to an Academic Review Board for further assessment. Students are not awarded academic credit hours for the comprehensive oral examination (and it will not appear on student transcripts), but the exam will be weighted as a 2-credit-hour course for the determination of Distinguished Graduates and Superior Graduates.³⁸

Preparing for and conducting comprehensive oral exams is a significant investment of faculty and student time that is spread across several months. Individual faculty members consume at least the equivalent of a week in preparation and another week in administering and grading the examinations. Mentoring a student who failed their first attempt can add significant faculty effort. The exam consists of a slate of questions that allow the students to show their achievement of the program learning outcomes. The questions are devised by a faculty committee, which incorporates input representing the expertise within each teaching department. The questions are revised each academic year to reflect changes in the curriculum but retain their focus at the program (not course) level.

Preparing for the Exams

The students begin preparing individually for the examination early in the 10-month program, connecting lesson and course learning outcomes with the higher program learning outcomes with coaching from their faculty team. The mere existence of the exam changes the dynamics of learning. Students and faculty are incentivized to accumulate knowledge and integrate it across courses to improve judgment. For example, a student who needs help understanding an important but intellectually challenging distinction being made in the seminar is incentivized to ask for clarification rather than remain silent until the class moves past the topic. The authors

³⁸ AY24 Resident Education Program Student Catalog, 33–34.

have, on multiple occasions, observed students seek clarification from peers and faculty in and out of the classroom precisely because they needed to be able to employ these concepts in an integrated fashion during comprehensive examinations. Absent this requirement, the workload and pace of the program would have led them to fail to understand the point, and because strategic studies, unlike math, are not visibly cumulative, they would have achieved a weaker outcome as a result, but it would not have necessarily been evident to an observer.

The comprehensive exam also changes faculty incentives. USAWC faculty are given broad latitude and academic freedom as long as their students accomplish specified learning outcomes. Knowing that outside faculty will assess their students on those outcomes helps keep an individual faculty member focused on the common outcome without overly constraining how they choose to conduct their seminar. They are incentivized to collaborate across the three departments represented on each seminar's faculty team to ensure that the students are not simply learning the material taught by each department based on its expertise but are developing holistically and making connections that allow them to translate learning into action—or at least talking about action. None of this is free in terms of student, faculty, and staff time, and the department chairs carefully consider the tradeoffs in the use of this time in coordination with the dean of the school.

Each year, the comprehensive exams committee develops five to six standard questions designed to allow the student to show integration of material across the curriculum while addressing real-world security issues. Students begin to prepare intensively in the eight weeks between the release of the examination questions and the exam's conduct. They must be able to deliver their answer to any of the slated questions orally, within about five minutes, using no notes. They craft their answers by referring to the PLOs and their course material, while consulting a rubric listing criteria for successful answers in *integrating course concepts*, *strategic thinking*, and *communication*. Each faculty team is authorized to conduct a single two-hour preparation session with their home seminar. No further faculty coaching is allowed, and the students collaborate to prepare and sharpen their answers.

Conducting the Exams

Most students sit for their comprehensive exams during a four-day window following the completion of the last core course. Students in two of the specialized programs (ASAP and NSPP) are not available to take their exams during this period because of academic travel, so they test on a single day several weeks later. To reduce possible faculty in-group bias, each student's exam is conducted by a faculty team other than the student's home seminar faculty. The registrar implements an additional safeguard against between-group bias by replicating the general contours of a base seminar in each test seminar, mixing Services and career specialties, and including several international officer students. Each faculty exam team also tests one or two students from a specialized program.

The comprehensive exams committee trains the faculty exam teams to offer a calibrated and common exam experience and assessment by employing a video, a faculty primer with examples, and a question-and-answer session. The dean of the school designates a lead faculty member within each three-person team to oversee that team's conduct of its exams.

During a 90-minute exam, students generally respond to three or four of the slated questions using their prepared answers. Much time is spent in dialogue generated by faculty follow-up questions inspired by the students' responses and the faculty's assessment of their demonstrated achievement of the PLOs.³⁹

Assessing Student Performance in the Exam

After about 60 minutes, the student is asked to leave the exam room while the faculty team deliberates on their performance using the rubric. The student returns to the room to receive immediate pass/fail feedback on their performance, again using the rubric. The faculty enters each student's letter grades from the rubric and personalized comments into the USAWC student information system, which is accessible to the student.

Students who do not meet standards for the exam receive support from their faculty and peers and may appeal the exam team's judgment

³⁹ AY24 Resident Education Program Student Catalog, 33–34.

through the dean to the deputy commandant. A third faculty team retests them after two weeks of studying for the same standardized questions; the second exam is graded as pass/fail. Only one student has failed the second exam since the 2020 implementation of PLO-based questions and faced an academic review board to determine their graduation status.⁴⁰

Evolution of USAWC Exams

The format and number of questions have evolved since the pilot in 2013 (table 1). The exam changed from focusing on course learning outcomes with questions developed at the department level to assessing integration across the program learning outcomes.

Two controversial but successful ideas emerged simultaneously in 2020 when the college developed six questions focused on the PLOs and released them to the students more than eight weeks before the examination. The authors discovered that the students became much more focused on what they thought was important to them and were much less likely to prepare either too narrowly or too generally. They also began to study collaboratively in groups to a greater degree, helping each other build a starting answer to rehearse, delivering that answer most effectively, and then practicing responding to unscripted follow-up questions. This practice also reduced student anxiety about the exams. The authors observed that students who performed poorly or failed the exam are more likely to have studied alone or done minimal group work than students who strongly meet or exceed the rubric criteria.

The authors assess that while the peer critiques are helpful, they matter less in setting up a student for success in the examination than having gone through the preparation to answer in front of one's peers and then having the experience of managing the stress of answering before a group of peers. Interestingly, even combat veterans with significant public speaking experience and previous advanced degrees commented to the authors that this is a stressful process and that group preparation was very important for their success in the examination. Most students the authors spoke to over the years after the examination felt it was a useful

⁴⁰ AY24 Resident Education Program Student Catalog, 34.

Table 1. Changes in the format of the comprehensive examination, by year

Academic year	Outcomes assessed	Development of questions	Questions released to students for preparation	Rubric	Notes
2013	Course Learning Outcomes (CLOs) for Core Courses			Pass or fail	Pilot year
2014	9 Institutional Learning Outcomes (ILOs)			Exceeds standards, meets standards, or fail	
2015	Joint Learning Areas (JLAs) from OPMEP			Aligned with USAWC course grading policy: exceeds standards (5 or 4), meets standards (3), needs improvement (2), or fail (1)	
2016	6 Program Learning Outcomes (PLOs)	30+ questions developed by the three departments	No		
2017					
2018	Experimentation with 2 departments using CLOs plus JLAs to develop questions and 1 department using CLOs while describing supported PLOs and JLAs for question development				
2019				New USAWC grading policy: A+ to F	
2020					
2021					Conducted on MS Teams because of COVID
2022	6 Program Learning Outcomes (PLOs)	6 program-level questions	Yes		
2023					
2024 and future					

Source: courtesy of the authors.

process to help them to integrate ideas from their time in the course and to prepare them for future high-stakes presentations.

Also of note, despite all students preparing with the same questions available, the faculty teams can easily distinguish the quality of the students' prepared answers and grade them with the aid of the rubric. Because the students clearly understand what they are being asked, faculty can make a confident judgment about their ability to think and communicate. Additionally, the students cannot perfectly anticipate the specific follow-up questions generated by the faculty's assessment of their initial response. Their impromptu responses and the iterative conversation allow faculty to make finer distinctions in the student's abilities. A student's lifetime work on thinking and communication skills matters as well as does their work during the course, but all students make a significant improvement during year. The best results are achieved when students have worked on these skills their whole life and then hone them during the year. However, some students arrive feeling that they are behind their peers on these skills and yet do very well on the examination as a result of intense and disciplined study, practice, and engagement with peers and faculty to develop their skills.

The USAWC comprehensive examinations were added to the curriculum as a major event was removed, a two-week Strategic Decision-Making Exercise (SDME) that previously ended the resident program core curriculum. A dedicated staff developed two simulated "worlds" complete with simulated crises experienced by the students in national security positions ranging from combatant command staff officers to senior leaders in Washington, DC. Faculty spent time preparing and playing roles in the two exercise worlds; students played for one week in each world in a different exercise role; and dedicated staff were needed to build and run the event. The learning achieved by each student was strongly correlated with the importance of their exercise role; however, many involved repetitious and not very challenging staff process actions, while other roles tested students' thinking and decision-making skills in high-profile situations. This inequity in skills development during the SDME—and a lack of rigorous assessment of student achievement of learning outcomes—led directly to the pilot of comprehensive examinations at USAWC.

The exams are certainly more useful for students, but they are also certainly more stressful than the SDME. It is difficult to say whether they actually consume more student time. The actual examination is only about 90 minutes, whereas the SDME occupied two weeks. The students certainly study a great deal for the examination, but that occurs in conjunction with their normal preparation for each lesson throughout the year. They learn to balance this preparation with researching and writing their research project and carrying out their individual learning plans to benefit themselves and their families.

Quantitative Analysis of Exam Data

The USAWC has developed a significant data set on student grades in their core courses and comprehensive exams. Using SPSS statistics, the authors conducted quantitative analysis of this data during 2021, 2022, and 2023 to address three questions that would be helpful to future work and would be of interest to other institutions. Although comprehensive exams have been underway since 2013, the data from these three years reflect student performance since shifting to allowing the students to prepare for known PLO-level questions. The full statistical results are available and summarized in this section.

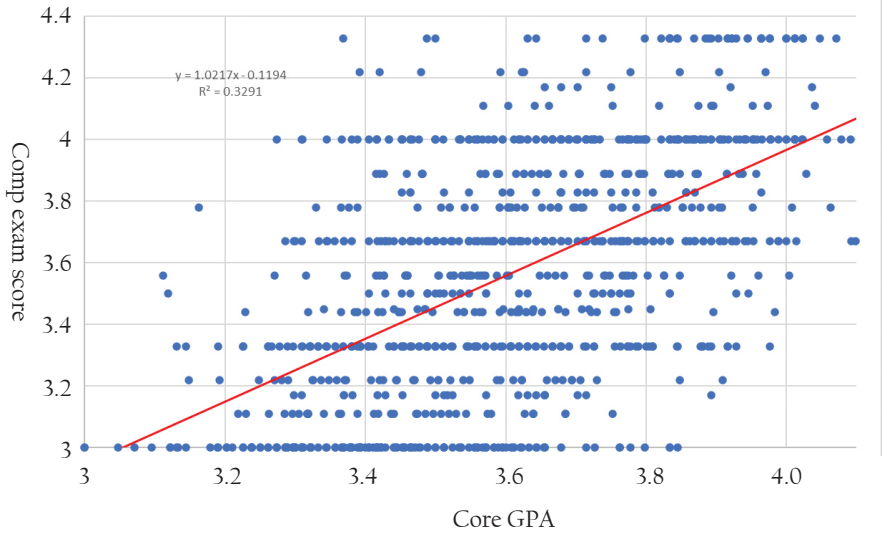
Research question 1: What is the relationship between course grades and comprehensive exam grades?

The first question the authors asked was the relationship between a student's grade point average in their core courses and their grade on the comprehensive exam. Figure 1 shows a high degree of correlation between a student's core course grade point average (GPA) and their examination grades. This result increases confidence that course learning outcomes are aligned with program learning outcomes.

Research question 2: What is the relative strength of PLOs 1–4 and 6 in influencing exam grades?

The second question addressed the relative strength of five of the six PLOs in influencing grades achieved by the students. Recall that the com-

Figure 1. Comprehensive exam grades and core course grade point average, 2021–23



Source: courtesy of the authors.

prehensive exam committee views the exam as most suitable for PLOs 1–4 and 6. The authors, therefore, developed two regression models, with the dependent variable being the comprehensive exam grade. In figure 2, five PLOs were treated as independent variables, and the authors found that the five PLOs explained 66.5 percent of the variance in comprehensive exam grades. PLO 6 accounts for 49 percent of that explanation, as can be seen from the adjusted R squared in figure 3.

PLO 6 for “strategic thinking and communication” has the most significant effect on comprehensive exam grades. Note that the exam is not designed to assess PLO 5 “collaboration.”

Research question 3A: Do exam grades provide evidence that students in specialized seminars experience better outcomes in achieving the PLOs than students in base seminars?

Since the students who volunteer for and are selected for the specialized

Figure 2. Five PLOs model

Model Summary				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.817 ^a	.667	.665	.342
a. Predictors: (Constant), PLO6, PLO1, PLO2, PLO3, PLO4				

Coefficients ^a					
Model		Unstandardized Coefficients		Standardized Coefficients	Sig.
		B	Std. Error	Beta	
1	(Constant)	.140	.089		1.562
	PLO1	.142	.024	.160	5.950
	PLO2	.131	.027	.135	4.839
	PLO3	.227	.030	.226	7.645
	PLO4	.145	.030	.143	4.772
	PLO6	.371	.029	.359	13.004

a. Dependent Variable: CompExam

Source: courtesy of the authors.

Figure 3. PLO 6 only model

Model Summary				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.700 ^a	.490	.490	.422
a. Predictors: (Constant), PLO6				

Coefficients ^a					
Model		Unstandardized Coefficients		Standardized Coefficients	Sig.
		B	Std. Error	Beta	
1	(Constant)	1.083	.093		11.621
	PLO6	.722	.027	.700	26.793

a. Dependent Variable: CompExam

Source: courtesy of the authors.

seminars are expected to be high performers in their academic work, do they perform differently on the comprehensive exam? The three “immersive specialization” seminars are:

- Advanced Strategic Art Program (ASAP)
- Carlisle Scholars Program (CSP)
- National Security Policy Program (NSPP)

Figures 4 and 5 showed that participation in a specialized seminar—without differentiating between the three specializations—is correlated with higher comprehensive exam grades. The statistical relationship is less significant than any of the PLO included in the model.

Research question 3B: Which specialized seminar is more significant in achieving the PLOs than students in base seminars?

The authors, therefore, tried another model that sought to make distinctions between the three programs by running a regression model (figure 6) with the dependent variable being the comprehensive exam grade and the independent variables being five of the five PLOs and membership in one of the three specialized seminars.

The results showed that two of the three programs had a statistically significant positive correlation with higher student comprehensive exam grades (ASAP and CSP). Further study is required to understand this differentiation between programs with volunteer student compositions and experienced faculty.

Limitations of the Study

Three limitations may affect the transferability of this study’s results to other PME institutions. First, a different mix of requirements, resources, faculty, and students may require a different approach. This study was limited to faculty and full-time students in the 10-month USAWC resident education program. In 2025, USAWC will expand comprehensive examinations to include the part-time students in the two-year USAWC distance education program. Second, the findings rely to a certain degree on professional judgment because the authors do not have a control

Figure 4. Model without specialized seminars

Model Summary				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.817 ^a	.667	.665	.342
a. Predictors: (Constant), PLO6, PLO1, PLO2, PLO3, PLO4				

Three “immersive specialization” seminars (who employ different curriculum sequencing and focus) take the same exam:

- Advanced Strategic Art Program (ASAP)
- Carlisle Scholars Program (CSP)
- National Security Policy Program (NSPP)

Source: courtesy of the authors.

Figure 5. Model with specialized seminars

Model Summary				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.819 ^a	.671	.669	.340
a. Predictors: (Constant), Special, PLO2, PLO1, PLO6, PLO3, PLO4				

Coefficients ^a					
Model		Unstandardized Coefficients		Standardized Coefficients	Sig.
		B	Std. Error	Beta	
1	(Constant)	.162	.089		.071
	PLO1	.139	.024	.157	<.001
	PLO2	.131	.027	.135	<.001
	PLO3	.224	.030	.224	<.001
	PLO4	.145	.030	.143	<.001
	PLO6	.365	.028	.354	<.001
	Special	.116	.039	.063	.003
a. Dependent Variable: CompExam					

Source: courtesy of the authors.

Figure 6. Model with five PLOs and three specialized seminars as independent variables

Model Summary				
Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.822 ^a	.676	.673	.338
a. Predictors: (Constant), NSPP, PLO4, CSP, ASAP, PLO1, PLO2, PLO6, PLO3				

Coefficients ^a						
		Unstandardized Coefficients		Standardized Coefficients		
Model		B	Std. Error	Beta	t	Sig.
1	(Constant)	.170	.089		1.915	.056
	PLO1	.142	.024	.160	6.009	<.001
	PLO2	.131	.027	.135	4.867	<.001
	PLO3	.224	.029	.224	7.631	<.001
	PLO4	.142	.030	.140	4.743	<.001
	PLO6	.363	.028	.352	12.843	<.001
	ASAP	.183	.066	.059	2.775	.006
	CSP	.223	.063	.074	3.513	<.001
	NSPP	-.048	.062	-.016	-.764	.445

a. Dependent Variable: CompExam

Source: courtesy of the authors.

group of full-time students who do not take comprehensive exams. They have sought to mitigate this issue by making conservative interpretations of the effects of regressions and by having the three authors constantly compare their interpretations of the results. Third, the implementation of these examinations consumed significant time and effort by academic leaders, staff, faculty, and, most importantly, the students.

The findings break with the trend toward negative observations in scholarly literature on comprehensive examinations. The authors assess this is at least in part because the institution, like CGSS, which also had a good experience described in the literature review, has much more substantial resources and incentives available to design, plan, prepare, execute, assess, and redesign these examinations on an annual basis than

does the average civilian graduate institution considered in the literature. It would be useful to find a way in a future study to further investigate the variables involved in making a comprehensive examination successful from the perspectives of various stakeholders, to further test observations, and to consider the impact of future innovations including the expansion of the test to include the nonresident faculty and students.

At USAWC, comprehensive examinations replaced an event that was arguably more time-consuming and less effective; it was a rare example for the three authors of an institutional change that improved both effectiveness and efficiency. They would not recommend simply adding something as time-consuming as a comprehensive examination without considering the impact on other uses of staff, faculty, and student time.

Lessons Learned in the Evolution of the Examination

The case study and regression analysis provide six key lessons.

1. Providing six standardized questions and the rubric to the students eight weeks prior to the test allowed them to prepare with a greater focus on the skills faculty expect graduates to face. This leads to an authentic assessment.
2. Training faculty on the rubric and conduct of follow-on dialogue is critical for calibrating the student experience and determining the exam grade.
3. The quality of the student's prepared response and the follow-on dialogue allows sufficient differentiation of performance for equitable assessment of the student's performance.
4. With informed student preparation using the current exam format, the first-time failure rate averages 5 percent yearly. The authors observed that these students are more likely than their peers who passed to have done little or no group preparation despite ample opportunities and encouragement from peers. The authors should neither expect a high failure rate given the highly competitive selection criteria to attend the course nor should they set

any passing percentage. The rubric establishes a standard of performance in line with program and higher guidance.

5. The significant correlation between students' course grades and comprehensive examination grades is consistent with the belief that both are effective assessments.
6. Participation in ASAP or CSP is a significant positive variable, while participation in the NSPP is not significant.

Conclusion

Collectively, these lessons suggest that an effective and efficient process of comprehensive examination has evolved at the USAWC. As PME for all ranks and Services continues to shift to OBME and authentic assessment, this case offers a useful model for consideration and adaptation to the individual circumstances of other programs. The authors conclude that comprehensive examination implementation and execution has been successful to date at the USAWC because, overall, it has provided more benefits than costs to the community of stakeholders.

Across the program, the authors have a high degree of confidence that the students are working more hours than they were a decade ago. Some portion of this is a result of the implementation of comprehensive examinations replacing other assessment events, but other requirements drive the bulk of this time increase. The more significant impact of the comprehensive examination has been to incentivize students to treat individual courses as a part of an integrated whole program of development rather than as a one-time event to be survived. This has made the entire program more meaningful as a developmental event for most students. In the period that follows the comprehensive examination, students' self-image often seems to change as they realize they are now capable of dealing competently with strategic issues that they may never have contemplated before this program.

For the faculty, the exams are useful to shape a motivated classroom. They incentivize every student to listen and learn even if they might not naturally be interested in the topic. Finally, they contribute to the freedom of the faculty in teaching core courses in a way that fits their style

and the particular needs of their students by providing a common reference point on which a faculty member can orient.

For the leadership and staff of the school and for outside stakeholders, including Army and Joint civilian and military leaders and academic accreditors, the comprehensive examination is a valuable component in a larger program of assessment and accountability. The examinations produce qualitative and quantitative data that supports the larger assessment of the outcomes rather than inputs of what is being taught in the program and how effectively the students are able to integrate and apply this to the type of problems and opportunities they may face in their future work. This is critical to assessing compliance with strategic guidance, with accreditation standards, and to the development of future guidance, standards, and initiatives.



CHAPTER 12 ●

Misconceptions among U.S. Air Force Academy Aviation Weather Cadets

An Exploratory and Validation Study
Using a Survey of Meteorology Concepts¹

*Lieutenant Colonel Daniel O'Keefe (USAF), PhD;
Major Louis J. Cascino (USAF); Matthew Mynk;
Senior Master Sergeant Nathaniel George (USAF);
Major Christopher Francis (USAF); Wilson González-Espada,
PhD; and Kimberly de La Harpe, PhD*

Introduction

Misconceptions are inaccurate conceptual understandings and cognitive structures rooted in students' thinking that deviate from the conception described by experts in their field, even after traditional instruction. Misconceptions influence how students receive and process new ideas and can interfere with their ability to form a meaningful understanding of complex concepts. Although researchers have explored misconceptions in several meteorology topics, a dearth of comprehensive, valid, and reliable questionnaires exist that can identify persistent misconceptions. Being well-informed, understanding the nature and impact of hazardous natural weather events, and implementing appropriate responses is essential for U.S. national security.

¹ This research was supported in part by the Department of Physics and Meteorology, United States Air Force Academy and through the Air Force Office of Scientific Research Summer Faculty Fellowship Program®, Contract Numbers FA8750-15-3-6003 and FA9550-15-0001. PA number USAFA-DF-2024-142. Approved for public release: distribution unlimited.

The purpose of this exploratory study was to develop the Survey of Meteorology Concepts (SMC) as a literature-based, psychometrically valid, and reliable meteorology questionnaire. Although this is still a work in progress and the sample size is only 98 cadets for the pre-surveys and 89 cadets for the post-surveys, the researchers discuss the validation strategies, which others can implement in their surveys and classroom-based assessments. These include (a) item response theory, including reliability coefficients, item difficulty, and item discrimination; (b) differential item functioning, a set of methods that can help identify disparities by gender, ethnicity, age, or education among cadets of the same ability; and (c) distractor analysis, a deep-analysis of all the alternatives in a multiple-choice question that, in association with a (d) certainty of response index prompt, can identify to what extent distractors are plausible, and which could represent an actual lack of knowledge of a topic or a misconception.

Higher education graduates with knowledge, skills, and abilities in science, technology, engineering, and mathematics (STEM) are essential for promoting innovation and economic growth in the United States.² However, universities are not producing enough STEM graduates to meet current and future demands in federal, national defense, private, and nonprofit sectors.³ This is a concerning situation for the

² Daman Huri, Noor Haslina, and Mageswary Karpudewan, "Evaluating the Effectiveness of Integrated STEM-Lab Activities in Improving Secondary School Students' Understanding of Electrolysis," *Chemistry Education Research and Practice* 20, no. 3 (2019): 495–508, <https://doi.org/10.1039/c9rp00021f>; and Louis S. Nadelson and Anne L. Seifert, "Integrated STEM Defined: Contexts, Challenges, and the Future," *Journal of Educational Research* 110, no. 3 (2017): 221–23, <https://doi.org/10.1080/00220671.2017.1289775>.

³ *State Department: Additional Actions Needed to Address IT Workforce Challenges* (Washington, DC: Government Accountability Office, 2022); *National Science Foundation: A Workforce Strategy and Evaluation of Results Could Improve Use of Rotating Scientists, Engineers, and Educators* (Washington, DC: Government Accountability Office, 2018); and Ronald Iammartino, John Bischoff, Christopher Willy, and Paul Shapiro, "Emergence in the U.S. Science, Technology, Engineering, and Mathematics (STEM) Workforce: An Agent-Based Model of Worker Attrition and Group Size in High-Density STEM Organizations," *Complex & Intelligent Systems* 2, no. 1 (2016): 23–34, <https://doi.org/10.1007/s40747-016-0015-7>.

U.S. Department of Defense, including the U.S. Air and Space Forces.⁴ In addition to the quantity of STEM graduates, their quality and ability to solve problems using accurate scientific information is an important concern.

Unfortunately, even those students who persist in STEM disciplines may hold a wide range of conceptual misconceptions. *Misconceptions* are defined as inaccurate conceptual understandings and cognitive structures that are rooted in students' thinking but deviate from the conception described by experts in their field, even after traditional instruction.⁵ Misconceptions influence how students receive and process new ideas and can interfere with their ability to form a meaningful understanding of complex concepts. As such, better understanding of misconceptions is relevant to STEM educators.⁶ For example, researchers have noted that students' physical geography misconceptions tend to develop preinstruction due to a combination of everyday experiences, myths, superstitions, movies, and social media misinformation, and are exceedingly difficult to unlearn.⁷ Research on minimizing misconceptions through a conceptual change model

⁴ *Improving the Air Force Scientific Discovery Mission: Leveraging Best Practices in Basic Research Management: A Workshop Report* (Washington, DC: National Academies Press, 2015), chap. 1; *Review of Specialized Degree-Granting Graduate Programs of the Department of Defense in STEM and Management* (Washington, DC: National Academies Press, 2014); *Assuring the U.S. Department of Defense a Strong Science, Technology, Engineering, and Mathematics (STEM) Workforce* (Washington, DC: National Academies Press, 2012); *Report of a Workshop on Science, Technology, Engineering, and Mathematics (STEM) Workforce Needs for the U.S. Department of Defense and the U.S. Defense Industrial Base* (Washington, DC: National Academies Press, 2011); and *Examination of the U.S. Air Force's Science, Technology, Engineering, and Mathematics (STEM) Workforce Needs in the Future and Its Strategy to Meet Those Needs* (Washington, DC: National Academies Press, 2010).

⁵ Philip M. Sadler and Gerhard Sonnert, "Understanding Misconceptions: Teaching and Learning in Middle School Physical Science," *American Educator* 40, no. 1 (2016): 26–32.

⁶ Phil Scott, Hilary Asoko, and John Leach, "Student Conceptions and Conceptual Learning in Science," in *Handbook of Research on Science Education*, ed. Sandra K. Abell, Ken Appleton, and Deborah Hanuscin (New York: Routledge, 2007), 31–56, <https://doi.org/10.4324/9780203824696>.

⁷ Jane Dove, "Reasons for Misconceptions in Physical Geography," *Geography* 101, no. 1 (2016): 47–53. <https://doi.org/10.1080/00167487.2016.12093983>.

became of interest to researchers in the 1980s and has developed theoretically since then.⁸

Occasionally, educators can contribute to the persistence of misconceptions by preparing tests and other assessments without an in-depth knowledge of educational measurement theory.⁹ They also assume that once instructors teach certain material, it is entirely understood by students, misjudging their extent and depth of understanding.¹⁰

Although researchers have explored misconceptions in meteorology topics such as clouds, vaporization and vapor pressure, climate change, air masses in the troposphere, and tornadoes, there is a dearth of comprehensive, valid, and reliable questionnaires that can identify persistent

⁸ George J. Posner et al., “Accommodation of a Scientific Conception: Toward a Theory of Conceptual Change,” *Science Education* 66, no. 2 (1982): 211–27, <https://doi.org/10.1002/sce.3730660207>; Panayiota Kendeou, Jason L. Braasch, and Ivar Bråten, “Optimizing Conditions for Learning: Situating Refutations in Epistemic Cognition,” *Journal of Experimental Education* 84, no. 2 (2015): 245–63, <https://doi.org/10.1080/00220973.2015.1027806>; and Stella Vosniadou, ed., *International Handbook of Research on Conceptual Change* (London: Routledge, 2013).

⁹ Douglas G. Bonett and Thomas A. Wright, “Cronbach’s Alpha Reliability: Interval Estimation, Hypothesis Testing, and Sample Size Planning,” *Journal of Organizational Behavior* 36, no. 1 (2014): 3–15, <https://doi.org/10.1002/job.1960>; Erin Ruel, “What Is Construct Validity?,” in *100 Questions (and Answers) About Survey Research* (Thousand Oaks, CA: Sage, 2019), 62, <https://doi.org/10.4135/9781506348803.n55>; Linda M. Crocker and James Algina, *Introduction to Classical and Modern Test Theory* (Mason, OH: Cengage Learning, 2008); Robert F. DeVellis, “Classical Test Theory,” *Medical Care* 44, no. 11 (2006): S50–S59, <https://doi.org/10.1097/01.mlr.0000245426.10853.30>; and Daniel M. Koretz, *Measuring Up: What Educational Testing Really Tells Us* (Cambridge, MA: Harvard University Press, 2009).

¹⁰ Casey E. Davenport, Christian S. Wohlwend, and Thomas L. Koehler, “Motivation for and Development of a Standardized Introductory Meteorology Assessment Exam,” *Bulletin of the American Meteorological Society* 96, no. 2 (2015): 305–12, <https://doi.org/10.1175/bams-d-13-00157.1>.

misconceptions.¹¹ The literature describes only one such instrument, the Fundamentals in Meteorology Inventory (FMI), which was partially developed from an internal assessment at the U.S. Air Force Academy (USAF) called the Meteorology Program Assessment Test.¹² However, the FMI lacks generality and broad applicability to introductory meteorology and pre-Service teachers taking Earth science classes for educators.

In the case of meteorology misconceptions, being a well-informed citizen who can understand the nature and impact of hazardous natural weather events and can implement appropriate responses is essential.¹³ Weather is inherent in military operations as well, and has been as far back as Sun Tzu's *Art of War*: "know the ground, know the weather; your victory will then be total."¹⁴ For the Air Force, studies have shown that weather-related crashes account for a significant proportion of aircraft

¹¹ Heidi Kroog and Maria Ruiz-Primo, "Are Clouds a Solid, Liquid, or Gas?," *Science and Children* 54, no. 8 (2017), https://doi.org/10.2505/4/sc17_054_08_41; Nurtaç Canpolat, Taçettin Pinarbaşı, and Mustafa Sözbilir, "Prospective Teachers' Misconceptions of Vaporization and Vapor Pressure," *Journal of Chemical Education* 83, no. 8 (2006): 1237, <https://doi.org/10.1021/ed083p1237>; Soyoung Choi et al., "Do Earth and Environmental Science Textbooks Promote Middle and High School Students' Conceptual Development about Climate Change?," *Bulletin of the American Meteorological Society* 91, no. 7 (2010): 889–98, <https://doi.org/10.1175/2009bams2625.1>; Vasiliki Papadimitriou and Paraskevi Londridou, "A Cross-age Study of Pupils' Conceptions Concerning the Movement of Air Masses in the Troposphere" (paper in proceedings of the IOSTE Symposium in Southern Europe, 29 April–2 May 2001, Cyprus); and Kelsey N. Ellis, Lisa Reyes Mason, and Kelly N. Gassert, "Public Understanding of Local Tornado Characteristics and Perceived Protection from Land-Surface Features in Tennessee, USA," *PLOS ONE* 14, no. 7 (2019), <https://doi.org/10.1371/journal.pone.0219897>.

¹² Casey E. Davenport and Adam J. French, "The Fundamentals in Meteorology Inventory: Validation of a Tool Assessing Basic Meteorological Conceptual Understanding," *Journal of Geoscience Education* 68, no. 2 (2019): 152–67, <https://doi.org/10.1080/10899995.2019.1629193>.

¹³ Rod Lane and Simon Catling, "Preservice Primary Teachers' Depth and Accuracy of Knowledge of Tropical Cyclones," *Journal of Geography* 115, no. 5 (2016): 198–211, <https://doi.org/10.1080/00221341.2016.1153133>.

¹⁴ Sun Tzu, *The Art of War* (New York, Dover Publications, 2002), chap. 10, para. 6.

accidents.¹⁵ At a minimum, pilots need to have as few meteorological misconceptions as possible and a high degree of weather literacy.

Developing a content knowledge survey requires the quantification of validity and reliability. *Validity* is the degree to which evidence exists to support the interpretation of test scores as accurately measuring the instructional objectives taught in class. *Reliability* is the degree to which an exam produces stable and consistent results.¹⁶ It also requires the implementation of item response theory (IRT), differential item functioning (DIF), and the certainty of response index (CRI).

IRT is a set of methods that richly describe the relationship between a latent trait (the construct to be assessed), the properties of individual items in the survey, and the participants' answers to individual items. Parameters of interest in IRT include: (a) Lee J. Cronbach's alpha or internal reliability of the test, which measures whether the items are consistently measuring the same latent variable; (b) item difficulty, a measure of the proportion of examinees who answered each question correctly; and (c) item discrimination, a measure of how well an exam question can differentiate between "upper performers," cadets who obtained the best scores on the questionnaire as a whole (70th percentile and above) and "lower performers," those who obtained the worse scores (30th percentile or

¹⁵ Joseph T. Coyne, Carryl L. Baldwin, and Kara A. Latorella, "Pilot Weather Assessment: Implications for Visual Flight Rules Flight into Instrument Meteorological Conditions," *International Journal of Aviation Psychology* 18, no. 2 (2008): 153–66, <https://doi.org/10.1080/10508410801926756>; David Ladwig, "Weather Hazards Training for United States Air Force (USAF) Flight Operations" (paper presented at the 25th AIAA Aerospace Sciences Meeting, 24 March–26 March 1987, Reno, NV), <https://doi.org/10.2514/6.1987-333>; and Mark Wiggins and David O'Hare, "Expertise in Aeronautical Weather-Related Decision Making: A Cross-Sectional Analysis of General Aviation Pilots," *Journal of Experimental Psychology: Applied* 1, no. 4 (1995): 305–20, <https://doi.org/10.1037/1076-898x.1.4.305>.

¹⁶ Krista D. Mattern and Brian F. Patterson, *Synthesis of Recent SAT Validity Findings: Trend Data over Time and Cohorts*, Research in Review 2014-1 (New York: College Board, 2014); and Emily J. Shaw, *An SAT Validity Primer* (New York: College Board, 2015).

below).¹⁷ Part of the IRT process also includes a close look at how all the distractors (incorrect alternatives) perform to confirm that they are perceived as plausible.

DIF is a set of methods that can help identify whether certain demographics (e.g., gender, ethnicity, age, or education) have different probabilities of correctly answering a given item on a multi-item survey or instrument after controlling for their overall scale scores.¹⁸ An example of this would be male cadets performing significantly better on an item compared with female cadets after matching members of each group by a certain ability level, such as overall test scores, or standardized measures, such as ACT scores.¹⁹ Specific to USAFA, the academic performance of cadets who were admitted after attending a preparatory school is a demographic of interest. Previous studies have shown that those who complete a preparatory school tend to struggle academi-

¹⁷ Douglas G. Bonett and Thomas A. Wright, "Cronbach's Alpha Reliability: Interval Estimation, Hypothesis Testing, and Sample Size Planning," *Journal of Organizational Behavior* 36, no. 1 (2014): 3–15, <https://doi.org/10.1002/job.1960>; Lee J. Cronbach, "Coefficient Alpha and the Internal Structure of Tests," *Psychometrika* 16, no. 3 (1951): 297–334, <https://doi.org/10.1007/bf02310555>; Laszlo A. Erdodi, "What Makes a Test Difficult?: Exploring the Effect of Item Content on Students' Performance," *Journal of Instructional Psychology* 39, no. 3–4 (2012): 171–76; and Frances M. Yang and Solon T. Kao, "Item Response Theory for Measurement Validity," *Shanghai Archives of Psychiatry* 26, no. 3 (2014): 171–77.

¹⁸ Georgina Guilera, Juana Gómez Benito, and M. Dolores Hidalgo, "Scientific Production on the Mantel-Haenszel Procedure as a Way of Detecting DIF," *Psicothema* 21, no. 3. (2009): 492–98; *Comparability of English and Spanish/English Mathematics Tests*, pt. 1, *Differential Item Functioning* (Salem, OR: Office of Assessment and Information Services, Oregon Department of Education, 2010); and Feride Ozyildirim Gumus, "Detecting DIF According to Gender and Liking Mathematics for Probability Problems given within/without Context," *International Online Journal of Educational Sciences* 11, no. 2 (2019): 118–30, <https://doi.org/10.15345/iojes.2019.02.008>.

¹⁹ Ángel M. Fidalgo et al., "Empirical Bayes versus Standard Mantel-Haenszel Statistics for Detecting Differential Item Functioning under Small Sample Conditions," *Journal of Experimental Education* 75, no. 4 (2007): 293–314, <https://doi.org/10.3200/jexe.75.4.293-316>; Wilson J. González-Espada, "Detecting Gender Bias through Test Item Analysis," *Physics Teacher* 47, no. 3 (2009): 175–79, <https://doi.org/10.1119/1.3081303>; Gary A. Morris et al., "Testing the Test: Item Response Curves and Test Quality," *American Journal of Physics* 74, no. 5 (2006): 449–53, <https://doi.org/10.1119/1.2174053>; and Steven J. Osterlind and Howard T. Everson, *Differential Item Functioning* 2d ed. (Thousand Oaks, CA: Sage, 2009), <https://doi.org/10.4135/9781412993913>.

Table 1. Summary and interpretation of the certainty of response index

	Low CRI (0–2)	High CRI (3–5)
Correct answer	Lack of knowledge (lucky guess)	Knowledge of correct concepts
Wrong answer	Lack of knowledge (unlucky guess)	Potential misconceptions

Note: individual CRI values are “0” for totally guessed answer, “1” almost a guess, “2” not sure, “3” sure, “4” almost certain, and “5” certain.

Source: Saleem Hasan, Diola Bagayoko, and Ella L Kelley, “Misconceptions and the Certainty of Response Index (CRI),” *Physics Education* 34, no. 5 (1999): 294–99, <https://doi.org/10.1088/0031-9120/34/5/304>.

cally and leave STEM fields at higher rates than those who do not.²⁰

The CRI is an option next to an item where respondents provide a self-assessed qualitative ranking of the degree of certainty in their ability to select and use the correct information to arrive at the answer, which the researchers converted into a numerical rank. If the CRI is low for an item, probably a degree of guesswork or a knowledge gap played significantly in the item response. High CRI values for an item indicates that test takers are noticeably confident in their knowledge when selecting an answer.²¹ A high average CRI for items incorrectly answered by many test takers may be an indication of misconceptions (table 1).

²⁰ Daniel S. O’Keefe, Wilson González-Espada, David Meier, “Beyond STEM Attrition: Quantifying the Flow of U.S. Air Force Academy Cadets between Academic Majors to Improve STEM Persistence,” *Journal of Military Learning* 7, no. 2 (2023): 3–25; and Daniel S. O’Keefe et al., “A Mixed Methods Analysis of STEM Major Attrition at the U.S. Air Force Academy,” *Journal of Military Learning* 6, no. 1 (2022): 15–38.

²¹ Saleem Hasan, Diola Bagayoko, and Ella L Kelley, “Misconceptions and the Certainty of Response Index (CRI),” *Physics Education* 34, no. 5 (1999): 294–99, <https://doi.org/10.1088/0031-9120/34/5/304>; and James M. Webb, William A. Stock, and Michael T. McCarthy, “The Effects of Feedback Timing on Learning Facts: The Role of Response Confidence,” *Contemporary Educational Psychology* 19, no. 3 (1994): 251–65, <https://doi.org/10.1006/ceps.1994.1020>.

Purpose and Research Questions

In Casey E. Davenport and Adam J. French's validation article, they noted that the participants found the FMI difficult; pre- and post-test scores for introductory meteorology students were 12.9/35 (36.9 percent) and 17.6/35 (50.3 percent), respectively.²² The researchers on this project are interested in developing the SMC as a literature-based, psychometrically valid, and reliable meteorology questionnaire that is general enough that it can be used both with USAFA cadets completing an introductory meteorology course and pre-Service teachers enrolled in an Earth and space science class at Morehead State University in Kentucky to identify persisting meteorology misconceptions after instruction (the pre-Service teacher results will be reported elsewhere).

The research objective of this exploratory study was to apply IRT, DIF, and CRI concepts to develop and validate the SMC, a locally constructed meteorology questionnaire that can help identify persistent cadet misconceptions in introductory meteorology and similar courses nationwide, making a significant contribution to the broader literature in Earth science education research. The research questions are:

- What questions in the SMC are, statistically, the best suited to measure weather-related misconceptions in terms of item response theory and the certainty of response index?
- What questions in the SMC are, statistically, the least likely to produce biased results in terms of differential item functioning?
- To what extent did USAFA cadets improve their meteorology knowledge and reduce their misconceptions after completing an introductory meteorology course?

Methodology

During the summer of 2023, 100 items were carefully constructed for the SMC based on common meteorology concepts and misconceptions identified in the literature. The researchers checked the items for content

²² Davenport and French, "The Fundamentals in Meteorology Inventory," 152–67.

accuracy and preliminarily validated them using a group of experts—faculty members with expertise in meteorology and geosciences. The experts ranked each question using this rubric:

- Not a priority: the question has meteorology knowledge that is not essential or does not address an important misconception.
- Okay, keep it: the question has important meteorology knowledge and addresses a misconception.
- Definitely keep: the question has essential meteorology knowledge and addresses a quite common misconception.

After removing the 10 questions rejected by the panel of experts, the 45 best-ranked questions were included in the first tested version of the SMC (SMC 1.0) that was piloted in the academic year 2023–24. Table 2 lists a summary of the questions and the number of alternatives. The remaining 45 questions will be included in the second version of the SMC (SMC 2.0), to be piloted in the academic year 2024–25. The top 45 questions (from SMC 1.0 and 2.0) shown to be reliable, valid, objective, and pinpoint misconceptions will, after additional editing and removing unattractive distractors, become the definitive version of the SMC. Additional data collection and validation with a sizable number of participants will be arranged.

Under the supervision of one of the researchers, Major Louis J. Cascino, cadets enrolled in Meteor 320 (Introduction to Meteorology and Aviation Weather) and completed the SMC 1.0 during the Fall 2023 and Spring 2024 semesters. Ninety-eight cadets completed it as a presurvey during lesson 2 (early in the semester), and 89 completed it as a post-survey during lesson 35 (near the end of the semester). Instructors used participation points to incentivize cadets and awarded points for completion and improvement on the post-survey. Cadets completed the survey during class time, and the instructor provided sufficient completion time.

Major Cascino compiled the SMC 1.0 pre- and post-results, including each item, alternatives, key, and how many cadets selected each alterna-

tive, as well as basic demographic information, such as gender, race, and preparatory school completion. Demographic information is required to perform the DIF analysis and identify potential item disparities. Since the researchers are reporting only two semesters worth of data, this study is exploratory, and its findings are preliminary in nature. However, the description of the SMC validation process can be of use by Scholarship of Teaching and Learning (SOTL) practitioners to adopt these techniques and increase the reliability and validity of other surveys.

Findings

Cronbach's Alpha

The internal reliability of a survey or instrument ranges from 0.0 to 1.0. Perfect internal reliability implies that all the questions are measuring the same thing, meteorology knowledge, in this case. In general, an internal consistency higher than about 0.7 is acceptable; an internal consistency higher than about 0.9 is excellent. The SMC 1.0 had a calculated Cronbach's value of 0.652. The output preliminarily flagged question 14 which, if removed, would increase the SMC 1.0's reliability to 0.671. As the researchers continue collecting data, they can revise or remove question 14 if necessary.

Item Difficulty

This metric ranges from 0.0 (when no cadets answered the item correctly) to 1.0 (when all cadets answered the item correctly). Items that are very easy (0.90–1.00) or very difficult (0.00–0.10) should be closely examined or modified for future use. For example, a question that almost no one got correct may indicate lack of knowledge or potential misconceptions. The research team found that the pre- and post-mean item difficulties were 0.403 ± 0.227 and 0.541 ± 0.250 , respectively. The analysis revealed that, for the presurvey, about 42 percent of the questions were difficult, 40 percent were moderately difficult, and about 18 percent of the questions were easy for cadets. In contrast, for the post-survey, about 22 percent of the questions were difficult or very difficult, 45 percent were moderately difficult, and about 33 percent of the questions were easy or very easy for

Table 2. Summary of SMC 1.0 questions and number of multiple-choice alternatives

Question number	Instructional objective (number of alternatives)
1	Rank gaseous chemical elements based on their abundance in the atmosphere (6).
2	Rank atmosphere layers based on their location with respect to the ground (6).
3	Graphically identify the rate of change in atmospheric pressure as a function of elevation (7).
4	Graphically identify the rate of change in atmospheric temperature as a function of elevation (7).
5	Graphically identify the daily variation in air temperature during a sunny, calm spring day (7).
6	Differentiate climate and weather (6).
7	Identify the type of weather, air mass rotation, and direction of Northern Hemisphere high atmospheric pressure regions (8).
8	Identify the type of weather, air mass rotation, and direction of Northern Hemisphere low atmospheric pressure regions (8).
9	Identify endothermic water phase changes (6).
10	State the change in volume and temperature of an air parcel flowing from a mountaintop to a valley (8).
11	Rank five substances based on their albedo (6).
12	Identify the cloud type associated with thunderstorms (6).
13	State in what month the Sun is at perihelion (7).
14	Identify the cause of differential solar radiation reaching the ground around midday and before sunset (6).
15	State how the summer and winter temperatures will change, if any, if the Earth's tilt decreases by 1.5° (6).
16	Rank four descriptions of sky conditions based on their cloud coverage (6).
17	Rank five wind barb diagrams based on wind speed (6).
18	Graphically identify cloud formation by convergence (7).
19	Graphically identify the relationship between raindrop diameter and terminal speed (7).
20	Graphically identify under what conditions a warm front produces rain, freezing rain, snow, and sleet (7).
21	Identify which of six wind barbs shows the direction of a northeasterly wind (7).
22	Identify the temperature characteristics and location of an El Niño weather pattern (8).

Source: courtesy of the authors.

Table 2. Summary of SMC 1.0 questions and number of multiple-choice alternatives (continued)

Question number	Instructional objective (number of alternatives)
23	Graphically identify the wind direction on an isobar map (7).
24	Graphically identify the prevailing wind direction in the Northern Hemisphere (7).
25	Graphically identify the weather symbol for a warm front (7).
26	Identify the most probable cause of climate change in the last 200 years (7).
27	Graphically identify the distribution of electric charges before lightning occurs (7).
28	Graphically identify wind motion and tornadoes based on a Doppler color image (6).
29	State the change in surface air pressure as a cold front pass (6).
30	Identify the Earth's region that experiences the largest annual range of average surface temperatures (6).
31	State what atmospheric conditions produce the coldest conditions at Earth's surface (6).
32	Identify the correct characteristics of the Coriolis force (5).
33	Identify the relationship between relative humidity, temperature, and air moisture (5).
34	State the proper relative humidity and condensation nuclei for cloud formation (6).
35	Identify the connection between thunderstorms and tornadoes (5).
36	State what causes the atmosphere to increase in temperature (6).
37	Identify what atmospheric features are made up of water vapor (8).
38	State the cause of the seasons (7).
39	Identify the air pressure and spinning direction characteristics of a hurricane (6).
40	State how warm and cold fronts interact when they meet (5).
41	State how frequently lightning strikes the same place (4).
42	Identify under what atmospheric temperature conditions water freezes and melts (6).
43	State what causes rainfall (6).
44	Identify the layers of the atmosphere that increase in temperature as the elevation increases (7).
45	State the causes of sea level rise associated with global warming (7).

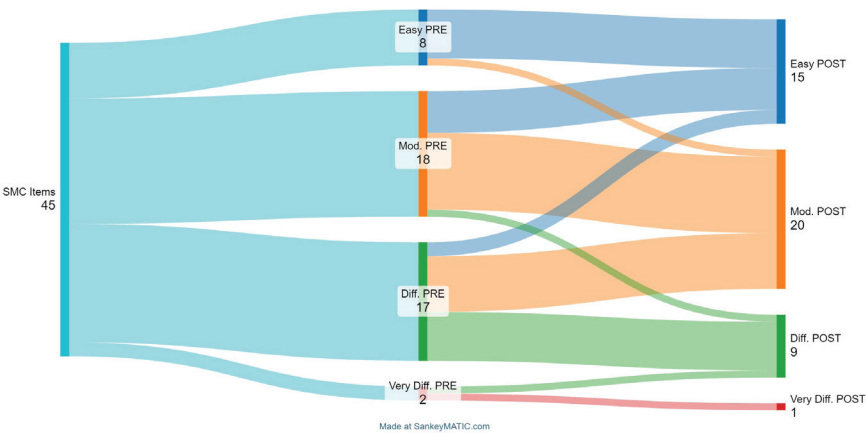
Source: courtesy of the authors.

Table 3. Number of SMC 1.0 questions categorized by item difficulty

Item difficulty	Presurvey items (%)	Post-survey items (%)
Very easy (1.00 to 0.90)	0	0
Easy (0.90 to 0.70)	8 (17.8%)	15 (33.3%)
Moderate (0.70 to 0.30)	18 (40.0%)	20 (44.4%)
Difficult (0.30 to 0.10)	17 (37.8%)	9 (20.0%)
Very difficult (0.10 to 0.00)	2 (4.4%)	1 (2.2%)

Note: percentages are calculated by columns.
Source: courtesy of the authors.

Figure 1. Sankey diagram flow chart showing item difficulty for pre- and post-survey



Source: courtesy of the authors.

cadets. Question 37 was perceived as very difficult in both surveys and should be further examined for revision or deletion. See table 3 and figure 1 for details.

Cadets found the SMC 1.0 questions easier during the post-survey. However, not all questions followed this trend to the same extent. After calculating each question's Richard R. Hake normalized gain (a coefficient that compares the difference between the pre- and post-survey while accounting for the pre-survey score), it was found that in 24 ques-

tions cadets had a relatively low gain (0–30 percent), in 15 questions the content gain was moderate (30–70 percent), and in two questions the gain was above 70 percent.²³ Notably, four questions had a negative gain—that is, the cadets found these post-survey questions more difficult than in the presurvey. A more comprehensive analysis of these questions, particularly if their CRI values were low (higher rates of guessing), could be used to narrow down the factors associated with the negative gains. For these four flagged questions, if this trend shows up again after adding more cadet results, the questions need to be closely examined and revised.

Item Discrimination

This metric ranges from -1.0 (when all lower performers answered the item correctly and all the upper performers answered the item incorrectly) to +1.0 (when all upper performers answered the item correctly and all the lower performers answered the item incorrectly). Items with discrimination less than 0.1 should be carefully examined, discarded, or modified for future use because it is uncommon for low scorers to perform much better on an item than high scorers.

The research team found that the pre- and post-mean item discriminations were 0.183 ± 0.124 and 0.252 ± 0.146 , respectively. The analysis revealed that, for the presurvey, about 77 percent of the questions had marginal or better discriminatory power, 16 percent had poor discrimination, and 7 percent had negative discrimination. After a semester of instruction, questions were better at discriminating between cadets of various levels of content knowledge; about 85 percent of the questions had marginal or better discriminatory power, 11 percent had poor discrimination and 4 percent had negative discrimination. Questions 3, 14, and 42 showed negative or poor discrimination on both surveys. Based on this analysis, seven post-survey questions should be further examined. See table 4 and figure 2 for details.

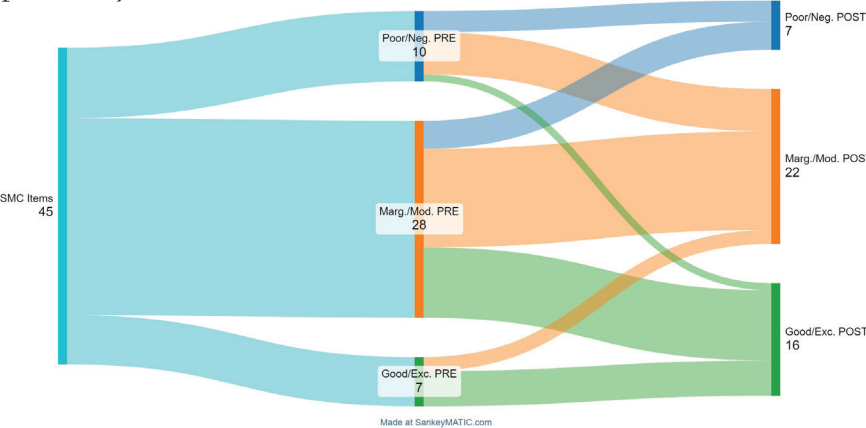
²³ Richard R. Hake, “Interactive-Engagement versus Traditional Methods: A Six-Thousand-Student Survey of Mechanics Test Data for Introductory Physics Courses,” *American Journal of Physics* 66, no. 1 (1998): 64–74, <https://doi.org/10.1119/1.18809>.

Table 4. Number of SMC 1.0 questions categorized by item discrimination

Item discrimination	Presurvey items (%)	Post-survey items (%)
Negative (-1.00 to 0)	3 (6.7%)	2 (4.4%)
Poor (0 to 0.10)	7 (15.6%)	5 (11.1%)
Marginal (0.10 to 0.20)	14 (31.1%)	10 (22.2%)
Moderate (0.20 to 0.30)	14 (31.1%)	12 (26.7%)
Good (0.30 to 0.40)	5 (11.1%)	10 (22.2%)
Excellent (0.40 to 1.00)	2 (4.4%)	6 (13.3%)

Note: percentages are calculated by columns.
Source: courtesy of the authors.

Figure 2. Sankey diagram flow chart showing item discrimination for pre- and post-survey



Source: courtesy of the authors.

Differential Item Functioning

For a large sample size, the Mantel-Haenszel (M-H) test, which is in the Chi-square family, creates a cross-tabulation table of correct and incorrect number of responses, for the previously matched set of examinees in both group memberships, over each of the several levels of ability, and indicates whether the common odds ratio of correct response across all

matched groups is similar to one (both groups performed similarly) or significantly different from unity (one group performed better than the other).²⁴

In this preliminary study, the sample sizes of female preparatory school completers and minority cadets limit the DIF analysis, and a formal M-H test is not possible. Figures 3 and 4 show an alternative analysis and summarize the results by percentile and three demographics of interest. The ratios of presurvey scores at the 50th percentile or above for males vs. females, prep school vs non-prep school completers, and white vs. minority cadets are roughly 1:1. In contrast, the ratios of post-survey scores are roughly 2:1 for prep school completers, revealing a potential disparity for this group.

Distractor Analysis

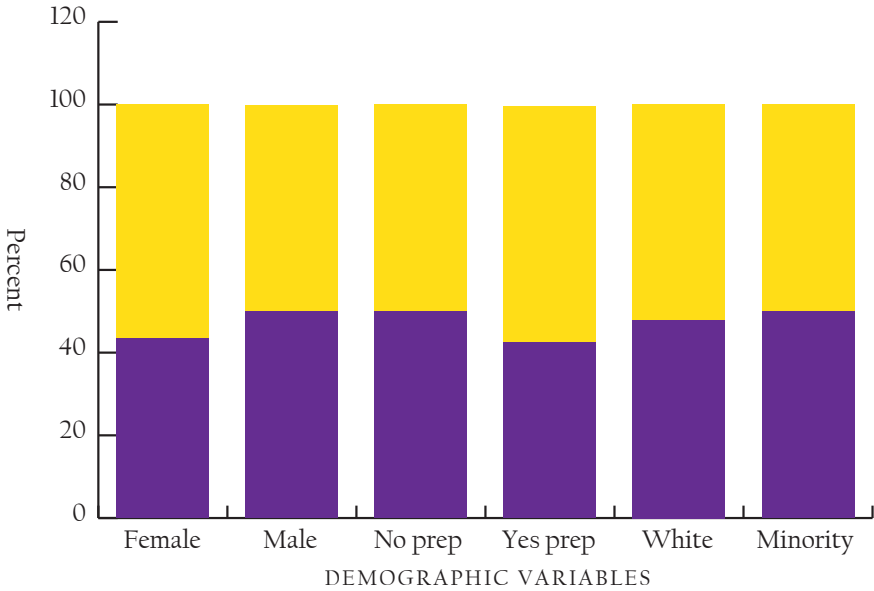
The researchers closely examined the responses for both the key and each distractor (incorrect alternative) for a given item. A good distractor should look reasonable and attractive for a test taker who does not know the correct answer but should be obviously incorrect for a test taker who knows the right answer. A nondistractor, an alternative that is rarely or never selected, should be closely examined or modified for future use. The traditional cutoff for nondistractors is less than 5 percent of the responses. Since it is difficult to foresee which distractors would be perceived as reasonable or immediately discarded, the researchers included more distractors than the 3–5 ideal number of distractors, based on the literature.²⁵

The analysis showed that about 58 percent of the presurvey items had two or more distractors that were rarely or never selected, increasing

²⁴ Ru Lu, Hongwen Guo, and Neil J. Dorans, *Robustness of Weighted Differential Item Functioning (DIF) Analysis: The Case of Mantel-Haenszel DIF Statistics*, Educational Testing Services Research Report 21-12 (Princeton, NJ: ETS, 2021), <https://doi.org/10.1002/ets2.12325>.

²⁵ Hongwen Guo, Jiyun Zu, Patrick Kyllonen, “A Simulation-Based Method for Finding the Optimal Number of Options for Multiple-Choice Items on a Test,” *ETS Research Report Series* no. 1 (2018), <https://doi.org/10.1002/ets2.12209>; M. Papenberg and Jochen Musch, “Of Small Beauties and Large Beasts: The Quality of Distractors on Multiple-Choice Tests Is More Important than Their Quantity,” *Applied Measurement in Education* 30, no. 4 (2017): 273–86, <https://doi.org/10.1080/08957347.2017.1353987>; and Thomas Haladyna, *Developing and Validating Multiple-Choice Test Items*, 3d ed. (Oxfordshire, UK: Routledge, 2015).

Figure 3. Percentage of cadets’ presurvey scores below median (purple) and above median (yellow) by demographic variables

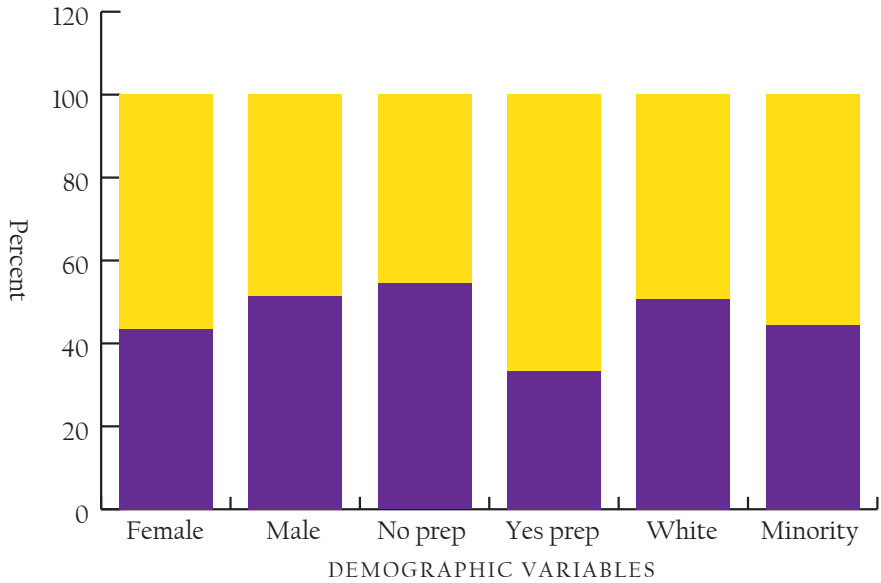


Source: courtesy of the authors.

to about 87 percent for the post-survey (table 5 and figure 5). This can be positive, as cadets are more knowledgeable after instruction and are less likely to be attracted to incorrect alternatives. However, additional nondistractors in one question decrease the theoretical chance of guessing the correct answer. For example, a cadet answering a question with 7 alternatives—3 of which are nondistractors—has a 14 percent chance of guessing the correct answer, compared to 25 percent if there were only 4 alternatives. Future analyses with an increased sample size and similar trends will aim for a final version of the SMC 1.0 with one key, three good distractors, and zero nondistractors.

Conversely, a distractor that is selected often or even more often than the correct alternative, such as an over-distractor, may indicate a misconception. Focusing on post-survey data, table 6 and figure 6 show that about 67 percent of the SMC 1.0 items included zero over-distractors. About 27 percent of the items had one or two over-distractors, includ-

Figure 4. Percentage of cadets' post-survey scores below median (purple) and above median (yellow) by demographic variables



Source: courtesy of the authors.

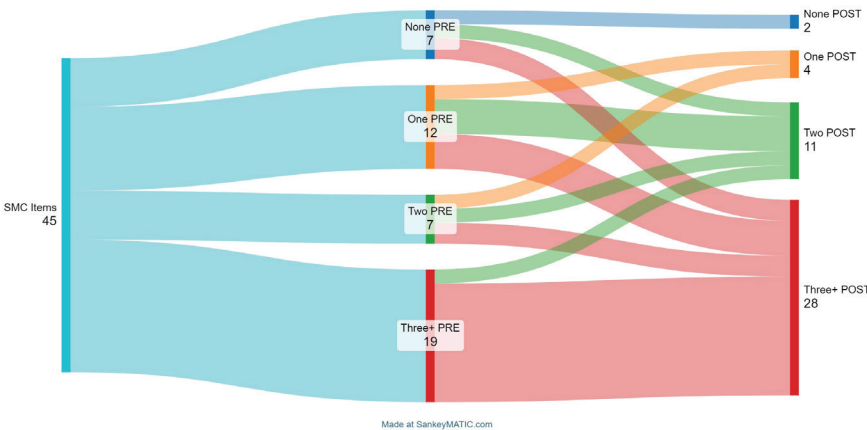
Table 5. Number of SMC 1.0 questions categorized by the number of nonselected alternatives (i.e., nondistractors).

Number of nondistractors in an item	Presurvey items (%)	Post-survey items (%)
None	7 (15.6%)	2 (4.4%)
One	12 (26.7%)	4 (8.9%)
Two	7 (15.6%)	11 (24.4%)
Three	10 (22.2%)	9 (20.0%)
Four	7 (15.6%)	10 (22.2%)
Five	1 (2.2%)	9 (20.0%)
Six	1 (2.2%)	0

Note: percentages are calculated by columns.

Source: courtesy of the authors.

Figure 5. Sankey diagram flow chart showing the number of nondistractors in an item for pre- and post-survey



Source: courtesy of the authors.

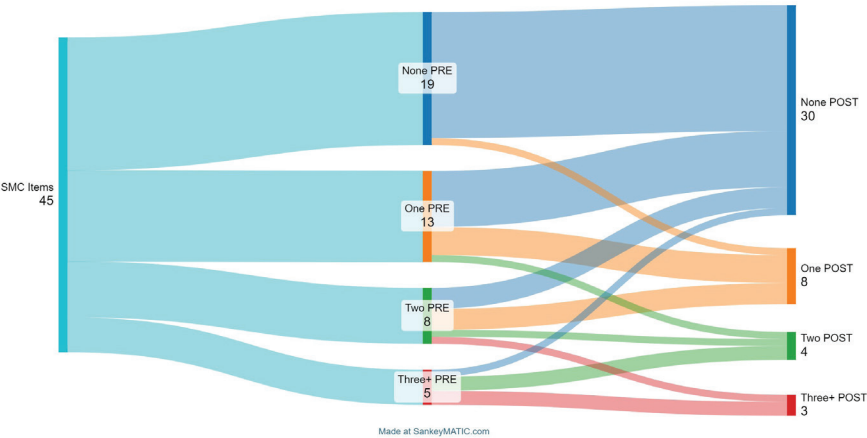
Table 6. Number of SMC 1.0 questions categorized by the number of multiselect-ed alternatives (i.e., over-distractors).

Number of over-distractors in an item	Presurvey items (%)	Post-survey items (%)
None	19 (42.2%)	30 (66.7%)
One	13 (28.9%)	7 (15.6%)
Two	8 (17.8%)	5 (11.1%)
Three	0	2 (4.4%)
Four	3 (6.7%)	1 (2.2%)
Five	1 (2.2%)	0
Six	1 (2.2%)	0

Note: percentages are calculated by columns.

Source: courtesy of the authors.

Figure 6. Sankey diagram flow chart showing the number of over-distractors in an item for pre- and post-survey



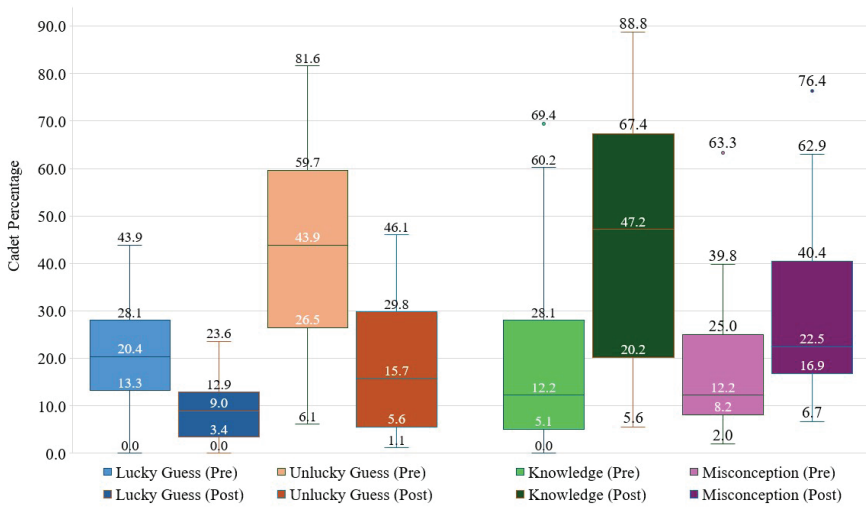
Source: courtesy of the authors.

ing questions 1, 19, 27, 31, 35, 36, and 41 with one over-distractor and questions 3, 11, 13, 34, and 42 with two over-distractors. Three or more over-distractors are usually seen with very difficult questions combined with significant guessing.

Certainty of Response Index (CRI)

This parameter ranges from zero (no knowledge; total guess) to five (complete confidence in the knowledge of the information required to arrive at the selected answer). Figure 7 shows that the number of guessed answers (i.e., CRI values of 0-2) decreases after instruction, as expected. Correct guesses decreased from an average of about 20 percent to about 9 percent, and incorrect guesses decreased from about 44 percent to about 16 percent. On the post-survey, questions with a substantial number of guesses should be rigorously evaluated to identify why cadets have low confidence in the content knowledge assessed by these items. This would be the case for questions 9, 19, 27, and 30 (20 percent or higher of lucky guesses), and questions 15, 19, 22, and 34 (40 percent or higher of unlucky guesses).

Figure 7. Box and whisker plot based on low CRI questions (lucky and unlucky guesses) and high CRI questions (correct knowledge and misconceptions) on the pre- and post-survey



Source: courtesy of the authors.

Figure 7 also shows that the number of correct answers where cadets showed high confidence in their content knowledge (i.e., CRI values of 3–5) increased after instruction. The average correct, high-confidence answers started at about 12 percent on the presurvey and finished at about 47 percent on the post-survey. With 80 percent or more of cadets answering correctly, cadets excelled in the content knowledge assessed by questions 6, 16, 17, 21, and 25. However, questions 3, 19, 29, and 37 had 10 percent or fewer correct responses.

The number of incorrect answers where cadets showed high confidence in their content knowledge, which would likely represent persistent misconceptions, increased after instruction. To an extent, this is an unexpected result. Ideally, the number of misconceptions should decrease after a semester of meteorology. However, one characteristic of misconceptions is that they are very persistent if not challenged in a direct way multiple times using conceptual change teaching strategies. It

is likely that cadets felt more confident of their content knowledge by the end of the semester but reverted to their old ideas in a timed testing situation. Another explanation could be a version of the Dunning Kruger effect—that is, after a semester meteorology course, cadets might consider themselves proficient with the material and have over-confidence when completing the SMC 1.0.²⁶ Meteorology instructors should critically examine the content knowledge covered in questions 1, 3, 29, 37, and 45 (i.e., post-survey questions answered incorrectly at confidence levels above 50 percent) and identify ways to challenge these misconceptions, both at USAFA and elsewhere.

Recommendations and Conclusions

Quantitative tools like IRT, DIF, and CRI are relatively easy to apply to multiple-choice tests and surveys given before and after instruction. In the case of the SMC 1.0, although the sample size impedes robust conclusions, the researchers were able to use cadet responses to identify questions that tended to be difficult, questions that may not be discriminating between cadets of different content knowledge levels, questions with potentially inadequate distractors, and questions that differentiated between actual content knowledge and persistent misconceptions. The data also showed that distinct groups responded to the SMC 1.0 differently, which raises critical disparity issues that must be further studied. The next step in the process of assessment evaluation is to identify potential causes of poor item performance, revise items, and continue the validation process.

One way to identify potential causes of poor item performance could be to map each question with the course's curriculum. This is important because not all meteorology textbooks cover the same material in the same way. Meteor 320 instructors noted, for instance, that more than 70 percent of the SMC 1.0 questions corresponded quite well with their lectures. Questions 6, 13, 19, 22, and 45 on the other hand, asked about topics not discussed in the course. Other questions have a different emphasis.

²⁶ J. Kruger and D. Dunning, "Unskilled and Unaware of It: How Difficulties in Recognizing One's Own Incompetence Lead to Inflated Self-assessments," *Journal of Personality and Social Psychology* 77, no. 6 (1999): 1121–34, <https://doi.org/10.1037//0022-3514.77.6.1121>.

Instructors stated that question 41 asked about how frequently lightning strikes the same place but explained that their teaching focus is on how lightning affects aircraft.

Another way to improve assessments is to evaluate whether the diagrams and the question wording are as precise as possible. For instance, Meteor 320 instructors noted that question 27 is phrased as if only the illustrated charge distribution on the ground and clouds is possible. They commented that lightning strikes from a negative ground to a positive cloud are possible but not common. Also, question 5 has several diagrams, and the instructor pointed out that two of them looked too similar to be separate distractors.

A third measure is to identify which questions required participants to apply two concepts in a single item. Ideally, each question should test a single concept. Otherwise, if the question is frequently missed, it is difficult to untangle which concepts the test takers are struggling with the most. Meteor 320 instructors suggested that question 18 should become two items, one dealing with cloud formation and another with surface convergence/divergence. Additionally, question 31 should be split, they argued, to deal separately with dew points and temperature changes at night.

After SMC questions are edited and revalidated, eventually a set of meteorology misconceptions will become clear. It is at this point that instructors need to critically examine their lectures and try innovative, hands-on teaching methods. Each misconception can be the starting point of an active learning demonstration, conceptual map, or laboratory experience, where data and analysis can help students challenge and modify their old ideas. From a broader perspective, given the prevalence of multiple-choice exams in postsecondary education courses, quantitative tools like IRT, DIF, and CRI are just as applicable to other disciplines, and their use is highly recommended.



CHAPTER 13

Data and Artificial Intelligence Literacy

Using Metrics and Experiential Learning
to Inform Andragogy¹

Kathleen Moore, PhD; and William Barry, PhD

Introduction

In recent years, the U.S. Army has sought to increase data literacy across its ranks of soldiers and officers as the Army transitions to what the U.S. Department of Defense calls a more “data-centric” organization where data is the primary strategic asset.² As the Army enterprise prepares for the next generation of technological advancements, the emergence of ChatGPT and other similar narrow-artificial intelligence (Narrow AI), large language models became available online worldwide, prompting an urgent dual call to add “AI literate” to the list of military personnel’s competencies.³ Narrow AI, also known as Weak AI or Artificial Narrow Intelligence (ANI), refers to AI systems that are designed to perform a specific task or a set of closely related tasks. Unlike general AI, which aims to replicate human intelligence and perform any intellectual task

¹ The content of this chapter was created, in part, as an amalgamation of the authors’ original materials from various courses, which was made possible by the use of an AI-assisted platform (Grammarly).

² Nathan Truckenbrod, *Data Literacy: A Necessary Pre-Condition for Making Data Driven Decisions* (Fort Eustis, VA: Army Training and Doctrine Command, 2023); and *Data Strategy*, 2020 (Washington, DC: Department of Defense, 2020).

³ Fred D. Davis Jr., “User Acceptance of Information Technology: System Characteristics, User Perceptions, and Behavioral Impacts,” *International Journal of Man-Machine Studies* 38 (1993): 475–87.

that a human being can, Narrow AI is limited in scope. It operates under a predefined set of rules and cannot exhibit the same level of understanding or adaptability as a human.⁴ In other words, Narrow AI focuses on handling a specific subset of tasks and is designed for tracking weather updates, generating data science reports by analyzing raw data, or playing games like poker or chess. Large language models, such as ChatGPT, are considered a form of Narrow AI.

At the time of this writing, the Army has not explicitly detailed the required skills and knowledge required across ranks and echelons for either endeavor. The U.S. Army War College, a professional military education (PME) institution, defines *data literacy* as the ability to read, write and communicate the value data in context to include understanding data types, the application of analytical methods including statistics and visualization, data governance, and ethics.⁵ Additionally, the college defines *AI literacy* as the ability of military professionals to critically evaluate AI technologies, communicate effectively with AI, and use AI as a tool in various contexts.⁶

In academic year 2023, the college sought to assess the data literacy of incoming students as one diagnostic measure of their overall capabilities as potential future strategic leaders. Consultation with other PMEs determined that no institution currently uses a data literacy assessment (DLA) tool. The college sought to identify a DLA that met the standards of peer review, reliability, and construct validity and began a systematic review of the academic literature that yielded few results. Further, the college determined the use of large language models (LLM) throughout the institution and incorporating the controlled use of such tools into its curriculum.

Analysis of the DLA literature revealed the challenges of teaching data literacy when the expected educators exhibit low rates of literacy them-

⁴ “Understanding Narrow AI: Definition, Capabilities, and Applications,” Deepai.org, accessed 27 September 2024.

⁵ From the syllabus of the Data and AI Informed Decision-Making course, academic year 2024, U.S. Army War College.

⁶ C. Anthony Pfaff et al., *Trusting AI: Integrating Artificial Intelligence into the Army's Professional Expert Knowledge* (Carlisle Barracks, PA: U.S. Army War College Press, 2023).

selves.⁷ Also, the instruction of data literacy is more effective when taught by trained educators.⁸ Further, using both applied problems as opposed to just testing for recognition of terms and concepts in assessments is a stronger indicator for literacy.⁹ A 2019 study examining multiple publicly available assessment tools concluded that none of the tests showed rigor in the form of reliability and construct validity and that many were a self-assessment of skills.¹⁰ Only one work developed a DLA that demonstrated reliability (albeit low), construct validity, and passed an expert panel review, but the test only uses concept recognition and not applied problem solving as its basis.¹¹

As the Army War College developed the DLA, professors simultaneously wanted to assess knowledge of and trust in machine learning and AI. While the college offered an elective on AI and human machine teaming, the introduction of open-source LLMs such as ChatGPT, and its use in the military domain, inspired conversations regarding not only the appropriate use of such technology but also whether to trust its use.

Similar to the DLA, a review of the academic literature determined that there is no peer-reviewed assessment pertaining to “AI Literacy,” which is the ability to know and understand machine learning and AI (ML/AI), to properly use and evaluate ML/AI, and to discern ethical issues of the technology.¹²

⁷ Bronwen Cowie and Beverly Cooper, *Exploring the Challenge of Developing Student Teacher Data Literacy* (London: Routledge, 2018).

⁸ Ailie McDowall et al., “Data Use as the Heart of Data Literacy: An Exploration of Pre-service Teachers’ Data Literacy Practices in a Teaching Performance Assessment,” *Asia-Pacific Journal of Teacher Education* 49, no. 5 (2021): 487–502, <https://doi.org/10.1080/1359866X.2020.1777529>.

⁹ Ying Cui et al., “Data Literacy Assessments: A Systematic Literature Review,” *Assessment in Education: Principles, Policy & Practice* 30, no. 1 (2023): 76–96, <https://doi.org/10.1080/0969594X.2023.2182737>.

¹⁰ Aneta Bonikowska, Claudia Sanmartin, and Marc Frenette, “Data Literacy: What It Is and How to Measure It in the Public Service,” Statistics Canada, 14 August 2019.

¹¹ Jeonghyun Kim et al., “Development and Validation of a Data Literacy Assessment Scale,” *Proceedings of the Association for Information Science and Technology* 60, no. 1 (October 2023): 620–24, <https://doi.org/10.1002/pra2.827>.

¹² Davy Tsz Kit Ng et al., “AI Literacy: Definition, Teaching, Evaluation and Ethical Issues,” *Proceedings of the Annual Meeting of the Association for Information Science and Technology* 58, no. 1 (2021), <https://doi.org/10.1002/pra2.487>.

The Army's Training and Doctrine Command (TRADOC) sets data literacy as a precondition for making data-driven decisions, and the fact that ML/AI is a hyper form of data analysis at its core means that data literacy must also be a precondition for AI literacy as well. This work documents the approach and steps taken by professors at the Army War College in developing a Data and AI Informed Decision-Making course to test this hypothesis for the academic year 2024.

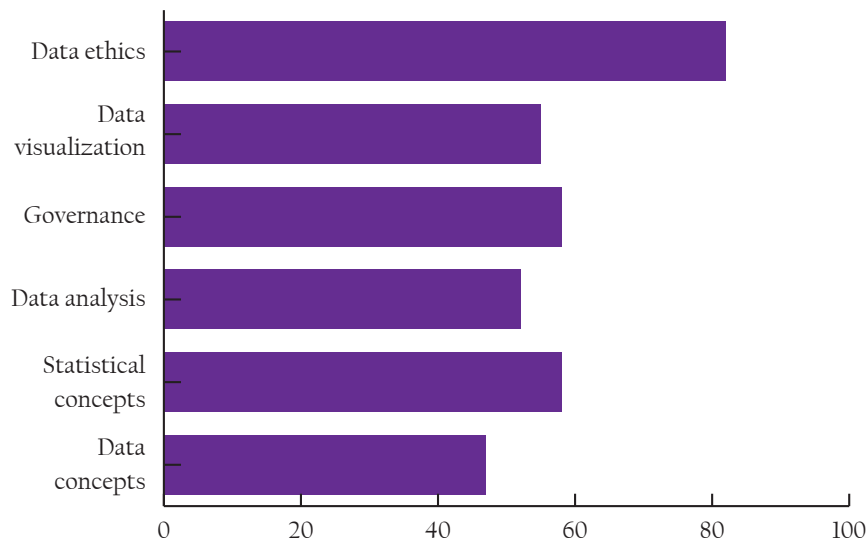
Data Literacy Assessment

In the absence of identifying a suitable assessment, the Army War College began the process of developing an internal DLA for its students based on the definition of literacy adopted by the institution. The first section of the test asked six questions in the areas of data concepts, statistical concepts, data analysis, data governance, data visualization, and data ethics. The second section asked for self-assessment in a variety of data skills. A pilot of the test occurred in Spring 2023 with faculty, staff, and students at the college. Professors revised questions for clarity before issuing the DLA in August 2023 on the incoming cohort of students. To gain more insight pertaining to the assessment quantitative results, all assessment tools allowed for respondents to provide qualitative feedback that allowed researchers to “measure and evaluate” and understand the phenomenon with more nuance.¹³

For that cohort, 354 incoming students took the DLA with a mean score of 58.33 out 100 percent. Initial results shown in figure 1 shows that students performed lowest on data concepts and highest on data ethics. From the internal demographics captured from test takers, students with STEM (science, technology, engineering, mathematics) degrees scored higher than social science and liberal arts majors. Students with doctorate degrees scored higher than those with master's or bachelor's degrees. Domestic students performed higher than international fellows, suggesting language barrier concerns. Also, students who took a data-centric class

¹³ Sae Schatz and J. J. Walcutt, “Modeling What Matters: AI and the Future of Defense Learning,” *Journal of Defense Modeling and Simulation: Applications, Methodology, Technology* 19, no. 2 (2022), <https://doi.org/10.1177/15485129221088718>.

Figure 1. Percent correct on the DLA by the 2024 cohort (N=354)



Source: courtesy of the authors.

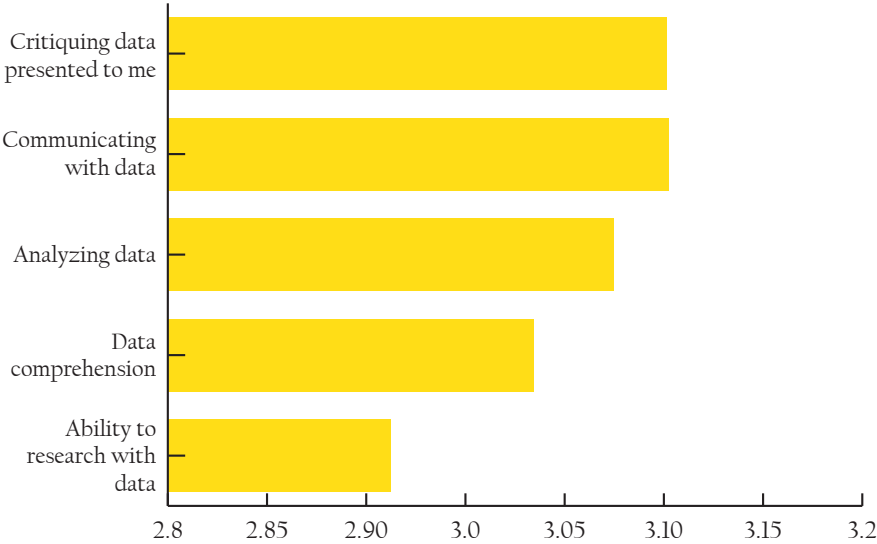
within the last 10 years, or who used data in their daily work scored higher overall.

In the affective portion of the DLA, students self-assessed skills qualitatively on a scale of 1 (low) to 5 (high) on their ability to perform research with data, comprehend data, analyze data, communicate with data, and critique data presented to them. Overall, the cohort rated itself 3.04 on average across the skill set (figure 2). The students rated themselves the lowest on the ability to research with data and general comprehend data and rated themselves higher on their ability to communicate with data and critiquing data presented to them in some manner.

A follow-on survey issued to 14 students taking the data literacy elective asked the respondents to expand on their answers from the DLA. From the survey, students reported three common themes from the course:

1. While I perform analysis and create charts all the time, I don't actually understand why I use some of the tools I use.

Figure 2. Qualitative self-assessment on data skills (N=354)



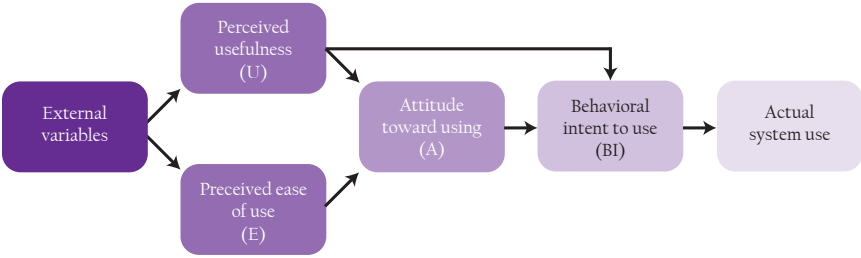
Source: courtesy of the authors.

- 2. I never knew much about data or technology. I find a lot of this confusing, and I don't know how the analysis works.
- 3. It's not that I don't trust the data or analysis per se as much as I don't trust myself to be the decision maker.¹⁴

Students expressed a need for a refresher-style course if not remedial instruction. Further, lack of trust in themselves reflects low confidence in their skills. The results of the assessment will help inform future interventions where students may take additional courses at the war college in the form of an elective or engage with online resources to help them develop their skills in areas where they require assistance. Professors used the results of this assessment to inform the redevelopment of an existing data literacy elective.

¹⁴ Comments extracted from follow-on survey issued to students of the Data and AI Informed Decision-Making elective, AY2024, U.S. Army War College.

Figure 3. Technology acceptance model



Source: Fred D. Davis Jr., “User Acceptance of Information Technology: System Characteristics, User Perceptions, and Behavioral Impacts,” *International Journal of Man-Machine Studies* 38 (1993), adapted by MCUP.

Machine Learning and AI Assessment

While developing the DLA, war college faculty also developed an assessment that gauged familiarity and use of the most common ML/AI products among faculty, staff, and students, their intent to use the technology, and their trust in the technology to perform Army-specific tasks.

Using the technology acceptance model (figure 3), respondents answered a battery of questions concerning their perceptions of the ease of use and overall utility of ML/AI products and future intention to use these products.¹⁵ The model focuses on two key aspects of technology acceptance by an individual: perceived ease of use and perceived personal utility. Both factors impact attitudes toward using a given technology, while only usefulness impacts behavioral intentions.

Trust for this assessment is distilled into three core concepts: ability, integrity, and benevolence.¹⁶ Ability speaks to the technological capability of performing as designed or promised. Integrity refers to the consistency and soundness of performance. Lastly, benevolence refers to good faith dealings, which in the world of technology is interpreted as transparency of how the technology is designed or works. Respondents answered ques-

¹⁵ Davis Jr., “User Acceptance of Information Technology,” 475–87.

¹⁶ Kenneth D. Alpern, “What Do We Want Trust to Be?: Some Distinctions of Trust,” *Business & Professional Ethics Journal* 16, nos. 1/3 (1997): 29–45.

tions concerning Army-specific tasks and assessed their trust in ML/AI performing those tasks on a Likert scale assessment of 1 to 5 (1 = no trust, and 5 = complete trust).

The assessment garnered 183 responses from faculty, staff, and students across the Army War College; 54 students completed the survey. From that number, 19 students were international fellows, meaning students from a foreign country, 38 of the students were from the general student body population, and 27 were students who participated in a year-long strategic futures seminar where ML/AI products were used extensively in their projects crafted for senior Army leaders.

The futures students unsurprisingly showed higher levels of familiarity and use of ML/AI products and possessed higher scores on ease of use and utility in those products than general students and international fellows (figure 4). Further, the futures students showed an overall higher level of trust in ML/AI technology than those of their peers on the same assessment.

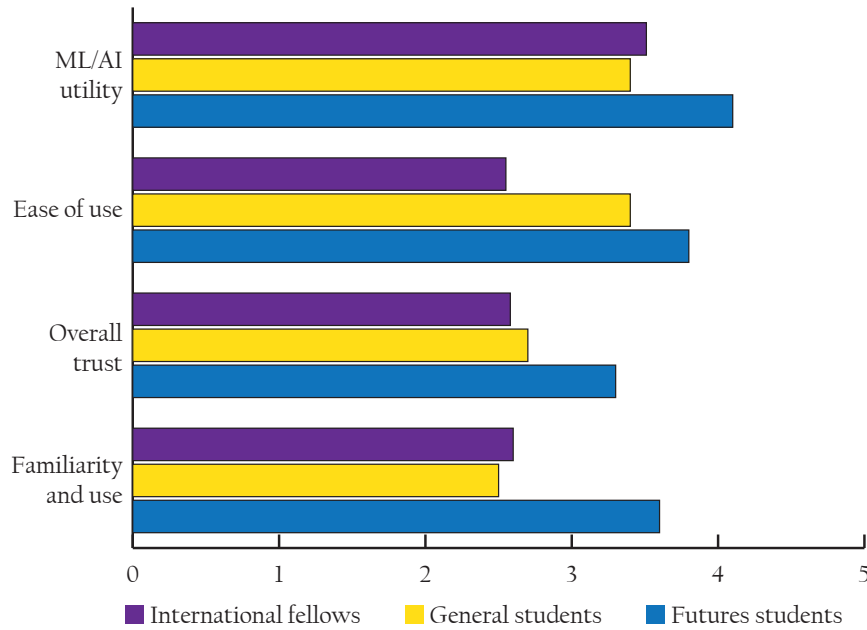
Additionally, the futures students also stated greater levels of intention (more than 90 percent) to use ML/AI technologies over those of their peers taking the same assessment (figure 5). Further, the same students expressed no intention to discontinue or be cautious of use of ML/AI technologies in the future, which differed greatly from their peers.

By examining the results, perceptions of utility and ease of use do not appear to be a barrier to technology acceptance of ML/AI, and while overall trust in the technology is higher among the futures students, their trust levels in ML/AI performing Army-specific tasking was generally on par with the whole of the survey population (table 1).

The 27 futures students were issued follow-up questions concerning the results and their class experience. Students largely considered the class experience to have impacted their responses on the ML/AI assessment (figure 6). When asked if they had registered low to zero trust on Army-specific tasking, they were then prompted to explain the response. Two themes emerged from 25 of the 27 respondents:

1. While we used the technology extensively in class, we were not instructed about how it works.

Figure 4. Student self-assessment concerning utility, ease of use, familiarity, and trust of ML/AI



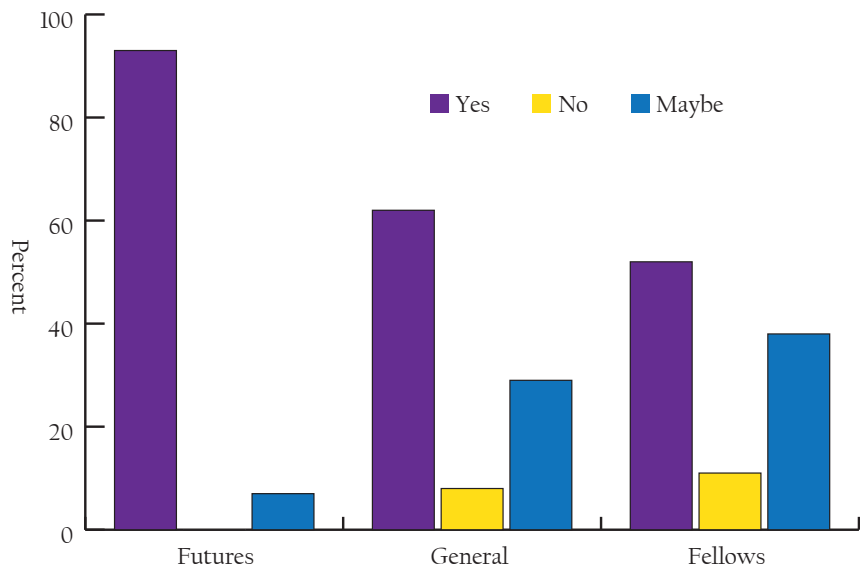
Source: courtesy of the authors.

- 2. I know the data process for many of these tasks and largely view them as problematic even without machine learning involved.¹⁷

When comparing the follow-on responses to the two assessments, the students’ comments show that it is not enough to interact with the tools, either in data analysis or ML/AI, but that students desire a deeper knowledge about how tools work. Also, students’ trust levels, either in their decision making or tools of analysis, are impacted by a fundamental lack of data knowledge literacy. A study has shown that positive emotions and technological self-efficacy (self-confidence) play a mediating role be-

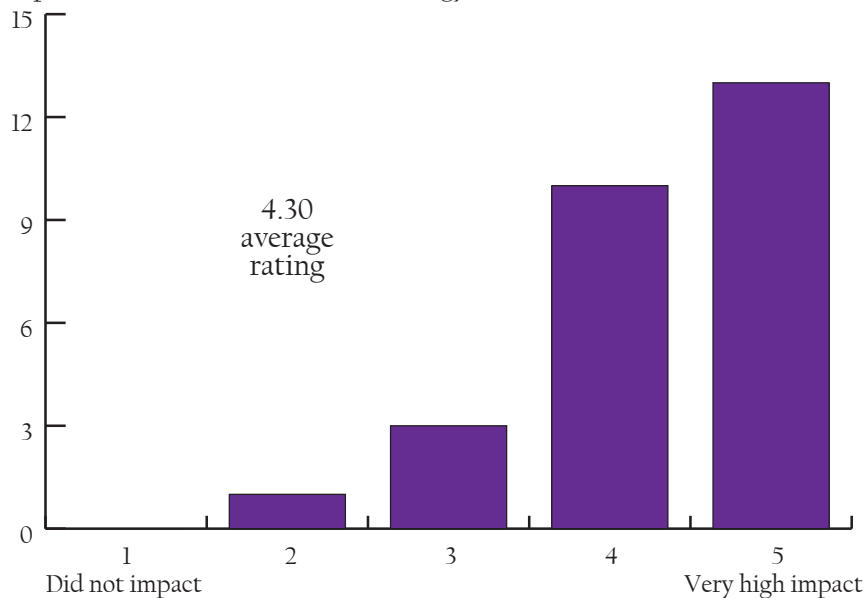
¹⁷ Comments extracted from a two question follow on survey to the ML/AI assessment.

Figure 5. Student self-assessment on intentions to use ML/AI in the future



Source: courtesy of the authors.

Figure 6. Futures students assessment of the impact the course had on their respective trust levels of ML/AI technology



Source: courtesy of the authors.

tween technology acceptance and self-directed learning.¹⁸ This suggests that self-confidence could potentially influence one's ability to trust and accept technology.

With these two analytical outcomes in mind, the data literacy elective for the war college underwent a complete rewrite in the summer preceding the academic year 2024 that merged the data literacy elective with the existing AI and human machine teaming elective into a single course.

Data and AI Informed Decision-Making

Overview

Data and AI informed decision making (DIDM) is a 27-contact hour elective offered during 9 individual course dates throughout the academic year. The course content addresses: data types, analysis (including qualitative and quantitative approaches), data-driven question development, data governance, data visualization, ML/AI, and data ethics. Since ML/AI is a hyper form of analysis and since data informs all analysis, each concept of the course is taught from the perspective of basic data literacy first, and then, how data informs ML/AI.

The instruction provided to students encompasses a comprehensive understanding of the fundamental tenets underlying data and AI literacy concepts, including their functions, limitations, practical applicability, operational mechanisms, and the contextual circumstances that may lead to their ineffectiveness. Students are given opportunities for multiple repetitions in class to practice using the concept with homework assignments that reinforce course content. Students then interact with the LLM that is primed with readings and activities for the course.

An LLM with a verbal and visual interface is used as a decision support tool where students reinforce the learning objectives of each class

¹⁸ Fuhai An et al., "Relationship between Technology Acceptance and Self-Directed Learning: Mediation Role of Positive Emotions and Technological Self-Efficacy," *Sustainability* 14, no. 16 (2022), <https://doi.org/10.3390/su141610390>.

Table 1. Assessing trust in ML/AI to perform Army-specific tasks comparing futures students’

Considering each task and activity below, what level of trust do you have for ML/AI to replace human participation?	Futures	General survey	Futures
	No trust	No trust	Low
Talent management	15%	43%	30%
Assessment	0	22%	15%
Recruitment	7%	17%	19%
Budgeting	0	13%	7%
Acquisition	0	19%	8%
Logistics	0	12%	4%
Force management	0	16%	30%
Intelligence analysis	0	28%	15%
Weapons effect analysis	0	12%	7%
Weapons release authority	37%	76%	29%

Source: courtesy of the authors.

through interaction with Technically Impossible Maneuver (TIM).¹⁹ For example, after a lesson on data terminology, students deconstruct a piece of data-driven research to identify types of data and analysis, and then interact with TIM, who either validates or corrects their responses. Data visualization is not used with TIM due to the limitations of the current framework.

TIM’s foundational framework is Open AI GPT3.5, a retrieval augmented generation design that uses bidirectional tuning. This design maximizes efficiency and use of tokens (units of text) when TIM makes inferences. Lastly, TIM is programmed with utilitarian ethics, which means that outputs aim for the greatest overall happiness or well-being and prescribes actions that maximize happiness and minimize suffering.

¹⁹ TIM is an intelligence augmentation system using large language models and context-specific data. See “Technically Impossible Maneuver (TIM),” Daxtron Laboratories, accessed 30 October 2024.

responses to the general survey population

General survey	Futures	General survey	Futures	General survey	Futures	General survey
Low	Moderate	Moderate	Moderately high	Moderately high	Extremely high	Extremely high
23%	37%	17%	7%	9%	11%	8%
4%	37%	40%	44%	18%	4%	16%
2%	33%	41%	37%	19%	4%	21%
10%	19%	27%	63%	34%	11%	16%
10%	35%	35%	50%	21%	7%	15%
12%	26%	19%	41%	24%	29%	33%
10%	44%	31%	19%	22%	7%	21%
28%	41%	21%	37%	10%	7%	13%
16%	30%	18%	52%	32%	11%	22%
8%	15%	8%	15%	5%	4%	3%

For example, when engaging with TIM for recommendations on balancing privacy and data use, students learn the importance of respecting individuals' privacy while still using data effectively, and TIM makes recommendations regarding anonymization techniques to protect privacy while still enabling useful data analysis.

For operational security, TIM does not touch open internet, and all data is vetted prior to import into TIM's memory. Uploads include the course syllabus, class slides, readings, Army regulations, doctrine plans, field manuals, technique publications, and the Law of Armed Conflict or international humanitarian law. Interactions with students are saved on a dedicated hard drive instead of the college network.

The intensive interactions and subsequent digital paper trail allow students to understand TIM's capabilities and deconstruct when and why TIM answers questions incorrectly. Students have multiple repetitions of practice with course concepts without the course instructor.

The following sections reflect the key areas where data literacy influ-

ences performance of ML/AI: data quality, data-driven questions as they pertain to prompt engineering, and governance including ethics.

Data Quality

Data quality plays a pivotal role in both general data and ML/AI contexts. The GIGO principle, the concept of “garbage in, garbage out,” emphasizes that poor input data leads to unreliable analytical results. In everyday contexts, data quality refers to the completeness, validity, uniqueness, consistency, conformity, and timeliness of information. It impacts decision making, business operations, and overall efficiency.

General Data Quality

Data sets ideally should be *complete*, meaning they contain all requisite elements and do not have empty data fields. The data set should also contain the correct data and be properly structured—in short, it should possess *validity*.

While version control may not always be possible, striving for a *unique* data set assures there will not be differences in analytical outputs. But should multiple versions of a data set exist, it should be certain that there is *consistency* and that there are no conflicts in outputs between sets or systems.

To analyze the data set, making sure that the set *conforms*, that the set adheres to preferred standards, ensures integrity and utility. And lastly, since data decays over *time*, recency in the data set is preferred.

Data Quality in ML/AI

Data quality is crucial for AI because it directly influences the performance, accuracy, and reliability of ML/AI models. High-quality data enables models to make more accurate predictions, fostering trust among users. Addressing biases in data is essential to avoid perpetuating biases in ML/AI-generated outputs. A diverse and representative dataset enhances a model’s ability to generalize well across different situations and contexts.

Data drift is a particular concern in ML/AI when characteristics and patterns of data change over time in unforeseen ways. The similarity be-

tween data drift and data consistency lies in the aspect of change. Both concepts involve changes in the data, although they occur in different contexts. Data drift focuses on the change in statistical properties over time, whereas data consistency focuses on ensuring that the data remains consistent and reliable despite potential changes.

In the context of machine learning, addressing data drift is crucial to maintain model performance and prevent degradation over time. It requires monitoring the model's performance, detecting drift, and adapting the model or retraining it with new data to account for the changes. Conversely, ensuring data consistency involves establishing data governance practices, data quality checks, and data validation processes to ensure that the data remains accurate and consistent across different sources and over time.

Internal issues with data quality can result in overfits, where a model memorizes training data instead of learning underlying patterns and generalizing to new examples. Contextual oversight is also of concern as ML/AI is not always able to fully account for broader data context. Lastly, while synthetic data mimics real word data and is frequently used to train models, it is not derived from actual observations, which can create issues of soundness in the model.

Recursive misinformation loops is another internal data quality problem where bad information is introduced and perpetuated, amplified through cycles of sharing, reinforcement, and validation. Lastly, external data quality issues can occur, such as when an adversary manipulates or modifies training data used in a machine learning model; so too can the mosaic effect, when unrelated pieces of information, when combined, reveal sensitive details about individuals. As students learn the fundamentals of data quality and advanced forms of analysis, they are better able to understand and foresee large-scale issues of data quality in ML/AI.

Developing More Robust Data Questions and Prompt Engineering

The purpose of a question is obviously to ask for and receive information, but questions provide context to research, which determines the type of data collected and analysis to be performed. The course is designed to

teach the fundamentals of SMART data-driven questions and the basics of prompt engineering when interacting with LLMs and ML/AI.

Data-Driven Questions

SMART data-driven questions are questions that are formulated using the SMART framework, which stands for specific, measurable, achievable, relevant, and time-bound. This framework is widely used in goal setting and project management to ensure that objectives are well-defined and actionable. When applied to data-driven questions, the framework helps in formulating clear and focused inquiries that can guide data analysis and decision-making processes effectively.

The first concept includes a specific question that is clear and well-defined, leaving no room for ambiguity. It addresses a particular aspect or problem and avoids being overly broad. For example, instead of asking, “How can we improve military readiness?,” a specific question would be, “What factors are contributing to the decline in performance for activity X impacting overall readiness?”

Next, a question should be measurable in that it can be quantified or evaluated objectively. It allows for the collection and analysis of relevant data to provide numerical or observable outcomes. For instance, instead of asking, “How effective was our mission?,” a measurable question would be, “What percentage of the mission objectives were achieved?”

Questions should also be achievable, or realistic and attainable within the available resources and constraints. It takes into consideration the feasibility of obtaining the required data and conducting the necessary analysis. For example, instead of asking, “What is the optimal number of companies required to secure the entire city?,” an achievable question would be, “What is the optimal number of companies required to secure the bridge?”

Next, questions should be relevant and directly related to the problem or objective at hand. It aligns with the overall goals and priorities of the organization or project. It ensures that the insights gained from data analysis are meaningful and useful for decision-making. For instance, instead of asking, “What is the average time it takes to complete a mission?,” a relevant question would be, “How does the average time it takes

to complete a mission change based on the terrain and the current weather conditions?”

Lastly, questions should be time-bound and have a timeframe or deadline associated with it. It sets a clear timeframe for data collection, analysis, and decision making. This component helps ensure that the question is actionable and time sensitive. For example, instead of asking, “What is the optimal number of companies required to secure the bridge?”, a time-bound question would be, “What is the optimal number of companies required to secure the bridge within the next 24 hours?”

By formulating data-driven questions using the SMART framework, students enhance the clarity, measurability, achievability, relevance, and time sensitivity of the inquiries, leading to more focused data analysis and actionable insights.

Prompt Engineering

The SMART framework and prompt engineering are related in the sense that both approaches aim to guide and structure the formulation of questions or prompts to achieve specific objectives. While the SMART framework is a general framework used in goal setting and project management, prompt engineering is a specific concept within the field of natural language processing (NLP) and language models.

Prompt engineering involves designing and refining the prompts or instructions given to language models to elicit desired responses. It focuses on finding the most effective and efficient ways to communicate with language models to obtain desired outputs. Prompt engineering techniques include careful phrasing, the use of context, specifying the desired format of the answer, and providing additional information or constraints. The SMART framework can be applied within the context of prompt engineering to create effective prompts.

Specific: A specific prompt clearly defines the task or objective for the language model. It precisely outlines the desired information or response to be generated. By providing specific instructions, the prompt can guide the model to produce the desired output.

Measurable: In prompt engineering, a measurable

prompt refers to a prompt that allows for objective evaluation of the generated output. It specifies the expected format, structure, or criteria for the response. By making the prompt measurable, it becomes possible to assess the model's performance against predefined metrics.

Achievable: An achievable prompt in prompt engineering takes into consideration the capabilities and limitations of the language model. It ensures that the prompt is within the model's scope and can be reasonably expected to produce a meaningful response. The prompt should be designed in a way that aligns with the model's capabilities.

Relevant: A relevant prompt is one that is directly related to the desired task or objective. It provides context and relevant information to guide the model's understanding and generate a response that is pertinent to the given prompt. A relevant prompt helps in eliciting focused and contextually appropriate answers.

Time-bound: In prompt engineering, a time-bound prompt may include time-related constraints or requirements. It can specify a desired response time or set a deadline for the model to generate the output. The time-bound aspect helps in managing expectations and ensuring timely responses.

By applying the principles of the SMART framework to prompt engineering, students can create well-defined, measurable, achievable, relevant, and time-bound prompts that effectively guide language models and improve the quality of generated responses.

Data Governance and Ethics

Data governance is crucial not only for data analysis but also for ML/AI because it ensures the quality, integrity, and responsible use of data throughout the entire lifecycle, including data acquisition, storage, processing, and decision making. For example, the assessments that form the basis of this research underwent internal review board (IRB) approval

prior to issuing the assessments to assure that the researchers protected the rights and welfare of human research subjects (students) recruited to participate. Informed consent was obtained prior to collection, and strict confidentiality and security of the data were maintained throughout the process under the auspices of the Army War College's IRB board. Lastly, all communication regarding the results is reviewed by the college's public affairs office prior to publication.

Data governance establishes standards and processes for data quality control, data validation, and data cleansing. High-quality and accurate data is essential for training ML/AI models, as the performance and reliability of these systems heavily rely on the quality of the input data. Data governance helps in maintaining data accuracy, consistency, and reliability, which in turn improves the performance and trustworthiness of models.

Good governance also ensures data security, privacy, and confidentiality where security measures and protocols protect sensitive data from unauthorized access, breaches, or misuse. This is particularly important in ML/AI, as these systems often process large volumes of data, including personal or sensitive information. Data governance helps in implementing security controls, data encryption, access controls, and data anonymization techniques to safeguard data privacy and prevent unauthorized use or disclosure of data.

No U.S.-based organization is exempt from regulatory compliance, and governance ensures compliance with relevant regulations and standards, such as data protection laws, industry-specific regulations, and internal policies. ML/AI applications must adhere to these regulations and standards to avoid legal and reputational risks. Governance provides mechanisms for data classification, data lineage, audit trails, and documentation, which are essential for demonstrating compliance and maintaining regulatory requirements.

Effective data governance also facilitates data integration and interoperability across different systems, data sources, and departments within an organization. ML/AI systems often rely on diverse data sources, and data integration is crucial for building comprehensive and accurate models. Governance helps in establishing data standards, data sharing

agreements, and data management practices that enable seamless data integration and interoperability, making it easier to leverage data for ML/AI purposes.

Data ethics, while often considered for purposes of morality or as a reflection of the values held by an organization, at its core, is the foundation for decision making.²⁰ Data ethics serves as a decision-making tool by providing a framework and guiding principles for making choices regarding data collection, storage, analysis, and usage. It helps organizations and individuals navigate the complex ethical challenges that arise in the context of data-driven technologies, including AI. Here is how data ethics relates to AI and decision-making.

Data ethics guides decision making in the development of ML/AI systems. It helps organizations consider the ethical implications of AI algorithms, data sets, and models. By incorporating data ethics into the development process, organizations can make informed decisions about the data they use, the biases they address, the potential impacts on individuals and society, and the fairness and transparency of their systems.

Additionally, data ethics promotes responsible data use in ML/AI applications. It helps organizations assess the implications of data collection, usage, and sharing. Decision-making guided by data ethics ensures that data is used in a manner that respects privacy, consent, and fairness. It involves considerations such as data minimization, purpose limitation, and ensuring consent-based data practices.

With regard to bias and fairness, data ethics helps address bias and promote fairness in ML/AI decision-making. It guides decision-making to mitigate biases that may be present in data sets, models, and algorithms. By considering data ethics, organizations can make decisions to identify, understand, and mitigate biases, ensuring that ML/AI systems do not discriminate against individuals or perpetuate unfair outcomes.

Data ethics encourages decision making that promotes transparency and explainability in ML/AI systems. It guides organizations to make decisions about providing explanations for ML/AI-generated outcomes, en-

²⁰ Eric Pliner, “A Framework for Leaders Facing Difficult Decisions,” *Harvard Business Review*, 13 October 2020.

abling individuals to understand how these systems make decisions that affect them. This transparency promotes accountability, trust, and allows affected parties to challenge decisions when necessary.

Lastly, data ethics prompts decision making that considers the broader societal impact of ML/AI. It encourages organizations to assess the potential consequences of these systems on individuals, communities, and society at large. By incorporating data ethics into decision-making, organizations can strive to ensure that their data systems align with societal values, mitigate harm, and contribute to the well-being of individuals and society.

In summary, data ethics serves as a decision-making tool by providing a framework to consider the ethical implications of data use and ML/AI development. It guides decisions related to system development, responsible data use, bias mitigation, fairness, transparency, and societal impact. By integrating data ethics into decision-making processes, students learn to make or guide informed choices concerning data systems that align with ethical principles and become future-focused on potential issues that arise as ML/AI systems mature and become integrated into future Army enterprise.

Limitations

Certain limitations exist in this work, as the population samples used in the assessments are purposeful and not generalizable to the greater military population. Additionally, both assessments are currently under review for statistical reliability and construct validity.

Finally, the students will not take the post-assessment on either data literacy or trust in ML/AI until the end of the academic year. Results of this analysis will likely inform revisions to the assessments themselves, but based on the qualitative data captured from respondents in the preassessment, this is unlikely to change the course design.

Conclusion

This chapter explored the interconnectedness between data literacy and ML/AI literacy, highlighting the significance of data literacy in informing AI literacy, and the role that self-confidence plays in trust and techno-

logical adoption. By examining the metrics and qualitative data captured from assessments, it emphasizes that a strong foundation in data literacy provides the necessary knowledge and skills to understand, evaluate, and effectively utilize ML/AI technologies. By comprehending the underlying data and its implications, individuals can make informed decisions, mitigate biases, promote fairness, and ensure responsible AI deployment.

At the completion of the 2024 academic year, students of the Data and AI Informed Decision-Making Course will complete a post-assessment assessing potential improvements in data literacy, the self-evaluation of skills, and changes in their self-assessment of trust in ML/AI.

Using qualitative data and grounded theory, an assessment of the approach used in this work will be conducted, and metrics from student performance feedback will inform refinements to future iterations of the course.



SELECT BIBLIOGRAPHY

- Abell, Sandra K., Ken Appleton, and Deborah Hanuscin, eds. *Handbook of Research on Science Education*. New York: Routledge, 2007.
- Adams, LtCol Owen, Reginald Shuford, Col Nathaniel Stone, LtCol Nicole Washington, and Col Dennis Weaver. *Human-Machine Teaming 2030–2040: Redefining the Continuum*. Carlisle, PA: U.S. Army War College, 2023.
- Advantage at Sea: Prevailing with Integrated All-Domain Naval Power, December 2020*. Washington, DC: Department of the Navy, 2020.
- Agresti, Alan, and Barbara Finlay. *Statistical Methods for the Social Sciences*, 4th ed. Upper Saddle River, NJ: Pearson Prentice Hall, 2009.
- Alencar, Eunice Maria Lima Soriano de, and Denise de Souza Fleith. “Criatividade na Educação Superior: Fatores Inibidores.” *Avaliação: Revista da Avaliação da Educação Superior (Campinas)* 15, no. 2 (2010): 201–19. <https://doi.org/10.1590/S1414-40772010000200011>.
- Amabile, Teresa M. “Social Psychology of Creativity: A Consensual Assessment Technique.” *Journal of Personality and Social Psychology* 43, no. 5 (1982): 997. <https://doi.org/10.1037/0022-3514.43.5.997>.
- Amabile, Teresa M., and Michael G. Pratt. “The Dynamic Componential Model of Creativity and Innovation in Organizations: Making Progress, Making Meaning.” *Research in Organizational Behavior* 36 (2016): 161. <https://doi.org/10.1016/j.riob.2016.10.001>.
- Anderson, Lorin W. “The Classroom Environment Study: Teaching for Learning.” *Comparative Education Review* 31, no. 1 (1987): 69–87. <https://doi.org/10.1086/446656>.

- Anderson, Wayne P., C. J. Krauskopf, M. E. Rogers, and G. W. Neal. "Reasons for Comprehensive Examinations: A Re-evaluation." *Teaching of Psychology* 11, no. 2 (1984): 78. https://doi.org/10.1207/s15328023top1102_4.
- Annual Suicide Report: Calendar Year 2020*. Washington, DC: Defense Suicide Prevention Office, Under Secretary of Defense for Personnel and Readiness, Department of Defense, 2021.
- Antinluoma, Markku, Liisa Ilomäki, and Auli Toom. "Practices of Professional Learning Communities." *Frontiers in Education* 6 (2021): 1–14. <https://doi.org/10.3389/feduc.2021.617613>.
- Archambault, Leanna, Heather Leary, and Kerry Rice. "Pillars of Online Pedagogy: A Framework for Teaching in Online Learning Environments." *Educational Psychologist* 57, no. 3 (2022): 178–91. <https://doi.org/10.1080/00461520.2022.2051513>.
- Armstrong, Patricia. *Bloom's Taxonomy*. Nashville, TN: Vanderbilt University Center for Teaching, 2010.
- Assuring the U.S. Department of Defense a Strong Science, Technology, Engineering, and Mathematics (STEM) Workforce*. Washington, DC: National Academies Press, 2012.
- Auten, Janet G., and Margaret M. Twigg. "Teaching and Learning SoTL: Preparing Future Faculty in a Pedagogy Course." *Teaching & Learning Inquiry* 3, no. 1 (2015): 3–13.
- Baker, David P. "Forward and Backward, Horizontal and Vertical: Transformation of Occupational Credentialing in the Schooled Society." *Research in Social Stratification and Mobility* 29, no. 1 (2011): 5–29. <https://doi.org/10.1016/j.rssm.2011.01.001>.
- Barab, Sasha, and Kurt Squire. "Design-Based Research: Putting a Stake in the Ground." *Journal of the Learning Sciences* 13, no. 1 (2004): 4. https://doi.org/10.1207/s15327809jls1301_1.
- Barnes, Adrienne E., and Young-Suk Kim. "Low-skilled Adult Readers Look Like Typically Developing Child Readers: A Comparison of Reading Skills and Eye Movement Behavior." *Reading and Writing: An Interdisciplinary Journal* 29, no. 9 (May 2016): 1889–914. <https://doi.org/10.1007/s11145-016-9657-5>.
- Barnett, Liz, David Brunne, Pal Maier, and Adam Warren. *Using Technology in Teaching and Learning*, 1st ed. London: Routledge, 1997.
- Barr, Robert B., and John Tagg. "From Teaching to Learning: A New Paradigm for Undergraduate Education." *Change: The Magazine of Higher Learning* 27, no. 6 (1995): 12–26. <https://doi.org/10.1080/00091383.1995.10544672>.
- Batanero, Carmen, G. Burrill, and C. Reading, eds. *Teaching Statistics in School Mathematics: Challenges for Teaching and Teacher Education*. Dordrecht: Springer, 2011.

- Beck, Dennis E., Leonel Morgado, and Patrick M. O'Shea. "Finding the Gaps about Uses of Immersive Learning Environments: A Survey of Surveys." *Journal of Universal Computer Science* 26, no. 8 (2020): 1043–73. <https://doi.org/10.3897/jucs.2020.055>.
- Beckman, Darrick, Jana Wardian, Tom J. Sauerwein, and Mark W. True. "Evaluation of an Interprofessional Continuing Professional Development Course on Comprehensive Diabetes Care: A Mixed-Methods Approach." *Journal of Evaluation in Clinical Practice* 25, no. 9 (October 2018): 149–50. <https://doi.org/10.1111/jep.13033>.
- Bennett, Sue, and Lori Lockyer. "Becoming an Online Teacher: Adapting to a Changed Environment for Teaching and Learning in Higher Education." *Educational Media International* 41, no. 3 (2004): 232–36. <https://doi.org/10.1080/09523980410001680842>.
- Berger, Gen David H. *Commandant's Planning Guidance: 38th Commandant of the Marine Corps*. Washington, DC: Headquarters Marine Corps, 2019.
- Bielaczyc, Katherine, and Allan Collins. "Learning Communities in Classrooms: A Reconceptualization of Educational Practice." In *Instructional Design Theories and Models*, vol. 2, *A New Paradigm of Instructional Theory*. Edited by Charles M. Reigeluth. Mahwah, NJ: Lawrence Erlbaum Associates, 1999.
- Bienkowski, Marie Ann, Mingyu Feng, and Barbara Means. *Enhancing Teaching and Learning Through Educational Data Mining and Learning Analytics: An Issue Brief*. Washington, DC: Office of Educational Technology, Department of Education, 2012.
- Blicblau, Aaron S., Megan Bruwer, and Kevin Dini. "Do Engineering Students Perceive that Different Learning and Teaching Modes Improve Their Referencing and Citation Skills?" *International Journal of Mechanical Engineering Education* 44, no. 1 (2015): 3–15. <https://doi.org/10.1177/0306419015624186>.
- Bloom, Benjamin S. *Taxonomy of Educational Objectives*. Boston, MA: Allyn and Bacon, 1956.
- Blunt, Adrian, and Baiyin Yang. "Factor Structure of the Adult Attitudes Toward Adult and Continuing Education Scale and Its Capacity to Predict Participation Behavior: Evidence for Adoption of a Revised Scale." *Adult Education Quarterly* 52, no. 4 (2002): 299–301.
- Bonett, Douglas G., and Thomas A. Wright. "Cronbach's Alpha Reliability: Interval Estimation, Hypothesis Testing, and Sample Size Planning." *Journal of Organizational Behavior* 36, no. 1 (2014): 3–15. <https://doi.org/10.1002/job.1960>.
- Bonikowska, Aneta, Claudia Sanmartin, and Marc Frenette. "Data Literacy: What It Is and How to Measure It in the Public Service." *Statistics Canada*, 14 August 2019.

- Boshier, Roger. "Psychometric Properties of the Alternative Form of the Education Participation Scale." *Adult Education Quarterly* 41, no. 3 (1991): 150. <https://doi.org/10.1177/0001848191041003002>.
- Boshuizen, Henny, Rainer Bromme and H. Gruber, eds. *Professional Learning: Gaps and Transitions on the Way from Novice to Expert*. New York: Springer, 2004.
- Boud, David, Rosemary Keogh, and David Walker, eds. *Reflection: Turning Experience into Learning*. New York: Routledge, 1985.
- Boyce, Lisa A., Stephen J. Zaccaro, and Michelle Zazanis Wisecarver. "Propensity for Self-Development of Leadership Attributes: Understanding, Predicting, and Supporting Performance of Leader Self-Development." *Leadership Quarterly* 21, no. 1 (2010): 160. <https://doi.org/10.1016/j.leaqua.2009.10.012>.
- Boyer, Ernest L. *Scholarship Reconsidered: Priorities of the Professoriate*. Princeton, NJ: Carnegie Foundation for the Advancement of Teaching, 1990.
- Bradley, Linda. "Peer-reviewing in an Intercultural Wiki Environment—Student Interaction and Reflections." *Computers and Composition* 34 (2014): 80–95. <https://doi.org/10.1016/j.compcom.2014.09.008>.
- Brennan, John, Steve Ryan, Marina Ranga, Simon Broek, Niccolo Durazzi, and Bregtje Kamphuis. *Study on Innovation in Higher Education: Final Report*. Luxembourg: Publications Office of the European Union, European Commission Directorate for Education and Training Study on Innovation in Higher Education, 2014.
- Briggs, Joseph, and Devesh Kodnani. *Global Economics Analyst: The Potentially Large Effects of Artificial Intelligence on Economic Growth*. New York: Goldman Sachs Publishing, 2023.
- Bronfenbrenner, Urie. *The Ecology of Human Development: Experiments by Nature and Design*. Cambridge, MA: Harvard University Press, 1979.
- Brown, Ann L. "Design Experiments: Theoretical and Methodological Challenges in Creating Complex Interventions in Classroom Settings." *Journal of the Learning Sciences* 2, no. 2 (1992): 141–78. https://doi.org/10.1207/s15327809jls0202_2.
- Brown, Benjamin D., Rolanda S. Horn, and Gwendolyn King. "The Effective Implementation of Professional Learning Communities." *Alabama Journal of Educational Leadership* 5 (August 2018): 53–59.
- Brown, John Seely, Allan Collins, and Paul Duguid. "Situated Cognition and the Culture of Learning." *Educational Researcher* 18, no. 1 (1989): 32–42. <https://doi.org/10.3102/0013189x018001032>.
- Brown, Peter C., Henry L. Roediger III, and Mark A. McDaniel. *Make It Stick: The Science of Successful Learning*. Cambridge, MA: Belknap Press, an imprint of Harvard University Press, 2014.

- Bryant, Susan F., and Andrew Harrison. *Finding Ender: Exploring the Intersections of Creativity, Innovation, and Talent Management in the U.S. Armed Forces*. Washington, DC: National Defense University Press, 2019.
- Burke, Susan K., and John T. Snead. "Faculty Opinions on the Use of Master's Degree End of Program Assessments." *Journal of Education for Library and Information Science* 55, no. 1 (2014).
- Capello, Alicia, and Janice Miller-Young. "Who Are We Citing and How?: A SOTL Citation Analysis?" *Teaching & Learning Inquiry* 8, no. 2 (2020): 2. <http://dx.doi.org/10.20343/teachlearninqu.8.2.2>.
- Carnevale, Anthony P., Leila J. Gainer, and Janice Villet. *Training in America: The Organization and Strategic Role of Training*. San Francisco, CA: Jossey-Bass, 1990.
- Chi, Micheline T. H., and Ruth Wylie. "The ICAP Framework: Linking Cognitive Engagement to Active Learning Outcomes." *Educational Psychologist* 49, no. 4 (2014): 219–43. <https://doi.org/10.1080/00461520.2014.965823>.
- Chick, Nancy, Lorelli Nowell, Bartłomiej A. Lenart. "The Scholarship of Teaching and Learning: A Scoping Review Protocol." *Teaching & Learning Inquiry* 7, no. 2 (2019). <http://dx.doi.org/10.20343/teachlearninqu.7.2.12>.
- Chua, Roy Y. J., Michael W. Morris, and Shira Mor. "Collaborating across Cultures: Cultural Metacognition and Affect-based Trust in Creative Collaboration." *Organizational Behavior and Human Decision Processes* 118, no. 2 (2012): 128. <https://dx.doi.org/10.1016/j.obhdp.2012.03.009>.
- Clark, Christopher M., Kate Olson, Ozge Hacifazlioglu, and David L. Carlson. "Community of Practice Among Faculty Team-Teaching Education Doctorate (Ed.D.) Students: A Reflective Study." *International Journal of Doctoral Studies* 16 (2021): 379–93. <https://doi.org/10.28945/4775>.
- Classen, Audra, Eva Horn, and Susan Palmer. "Needs of Military Families: Family and Educator Perspective." *Journal of Early Intervention* 41, no. 3 (2019): 246–47. <https://doi.org/10.1177/1053815119847235>.
- Coe, C. D. "Scaffolded Writing as a Tool for Critical Thinking: Teaching Beginning Students How to Write Arguments." *Teaching Philosophy* 34, no. 1 (2011): 33–50. <https://doi.org/10.5840/teachphil20113413>.
- Collins, Allan, Diana Joseph, and Katerine Bielaczyc. "Design Research: Theoretical and Methodological Issues." *Journal of the Learning Sciences* 13, no. 1 (2004): 15–42. https://doi.org/10.1207/s15327809jls1301_2.
- Combe, P. C., II. "Educational Wargaming: Design and Implementation into Professional Military Education." *Journal of Advanced Military Studies* 12, no. 2 (2021): 115–38. <https://doi.org/10.21140/mcu.20211202004>.
- Connors, Robert J., and Andrea A. Lunsford. "Frequency of Formal Errors in Current College Writing, or Ma and Pa Kettle Do Research." *College Compos-*

- sition and Communication 39, no. 4 (1988): 395–409. <https://doi.org/10.2307/357695>.
- Creative Thinking VALUE Rubric. Washington, DC: American Association of Colleges and Universities, 2009.
- Creswell, John W. *Educational Research: Planning, Conducting, and Evaluating Quantitative and Qualitative Research*. New York: Pearson Education, 2005.
- Creswell, John W., and Vicki L. Plano Clark. *Designing and Conducting Mixed Methods Research*, 2d ed. Los Angeles, CA: Sage Publications, 2011.
- Creswell, John W., and Cheryl N. Poth. *Qualitative Inquiry and Research Design: Choosing Among Five Approaches*. Thousand Oaks, CA: Sage, 2018.
- Crews, Tena B., Kelly Wilkinson, and Jason K. Neill. “Principles for Good Practice in Undergraduate Education: Effective Online Course Design to Assist Students’ Success.” *MERLOT Journal of Online Learning and Teaching* 11, no. 1 (March 2015): 87–103.
- Crocker, Linda M., and James Algina. *Introduction to Classical and Modern Test Theory*. Mason, OH: Cengage Learning, 2008.
- Cruz, Laura. “Converging Spaces?: Re-Imagining the Scholarship of Teaching and Learning as the Subject of Transdisciplinary Research.” *Transformative Dialogues: Teaching and Learning Journal* 17, no. 1 (Spring 2024). <https://doi.org/10.26209/td2024voll7iss11823>.
- Cui, Ying, Fu Chen, Alina Lutsyk, Jacqueline P. Leighton, and Maria Cutumisu. “Data Literacy Assessments: A Systematic Literature Review.” *Assessment in Education: Principles, Policy & Practice* 30, no. 1 (2023): 76–96. <https://doi.org/10.1080/0969594X.2023.2182737>.
- Cushman, Ellen. “The Rhetorician as an Agent of Social Change.” *College Composition and Communication* 47, no. 1 (1996): 7–28. <https://doi.org/10.2307/358271>.
- Czerkawski, Betul C., and Eugene W. Lyman III. “An Instructional Design Framework for Fostering Student Engagement in Online Learning Environments.” *TechTrends* 60 (2016): 532–39. <https://doi.org/10.1007/s11528-016-0110-z>.
- Darby, Flower, with James M. Lang. *Small Teaching Online: Applying Learning Science in Online Classes*. San Francisco, CA: Jossey-Bass, 2019.
- Data, Analytics, and Artificial Intelligence Adoption Strategy: Accelerating Decision Advantage*. Washington, DC: Department of Defense, 2023.
- Data Strategy*, 2020. Washington, DC: Department of Defense, 2020.
- Davis Jr., Fred D. “User Acceptance of Information Technology: System Characteristics, User Perceptions, and Behavioral Impacts.” *International Journal of Man-Machine Studies* 38 (1993): 475–87.
- Davis, Steven B., and John M. Hinck. “Assessing Instructor Developmental Stages

- in an Accelerated Schedule: A Guided Reflection.” *Journal of Faculty Development* 38, no. 3 (2024).
- De Freitas, Sara, Genaro Rebolledo-Mendez, Fotis Liarokapis, George Magoulas, and Alexandra Poulouvassilis. “Learning as Immersive Experiences: Using the Four-Dimensional Framework for Designing and Evaluating Immersive Learning Experiences in a Virtual World.” *British Journal of Educational Technology* 41, no. 1 (2010): 69–85. <https://doi.org/10.1111/j.1467-8535.2009.01024.x>.
- Department of Defense Instruction (DODI) 1322.35, Volume 1. *Military Education: Program Management and Administration*. Washington, DC: Department of Defense, 26 April 2022.
- Desimone, Laura M., and Kerstin Carlson Le Floch. “Are We Asking the Right Questions?: Using Cognitive Interviews to Improve Surveys in Education Research.” *Educational Evaluation and Policy Analysis* 26, no. 1 (2004): 1. <https://doi.org/10.3102/01623737026001001>.
- DeVellis, Robert F. “Classical Test Theory.” *Medical Care* 44, no. 11 (2006): S50–S59. <https://doi.org/10.1097/01.mlr.0000245426.10853.30>.
- Developing Enlisted Leaders for Tomorrow’s Wars: Our Shared Vision for Enlisted Professional Military Education & Talent Management*. Washington, DC: Joint Chiefs of Staff, 2021.
- Developing Today’s Joint Officers for Tomorrow’s Ways of War: The Joint Chiefs of Staff Vision and Guidance for Professional Military Education & Talent Management*. Washington, DC: Joint Chiefs of Staff, 2020.
- Developments in Business Simulation and Experiential Learning: Proceedings of the Annual ABSEL Conference*, vol. 50. Fort Leavenworth, KS: Army University Press, 2023.
- Devi, Barkha, Bidita Khandelwal, and Mridula Das. “Application of Bandura’s Social Cognitive Theory in the Technology Enhanced, Blended Learning Environment.” *International Journal of Applied Research* 3, no. 1 (2017).
- Dewey, John. *Experience and Education*. Indianapolis, IN: Kappa Delta Pi, 1938; New York, NY: Free Press, 1997.
- Doo, Min Young, and Jungwon Kim. “The Relationship between Learning Engagement and Learning Outcomes in Online Learning in Higher Education: A Meta-analysis Study.” *Distance Education* 45, no. 1 (2024): 60–82. <https://doi.org/10.1080/01587919.2024.2303484>.
- Dove, Jane. “Reasons for Misconceptions in Physical Geography.” *Geography* 101, no. 1 (2016): 47–53. <https://doi.org/10.1080/00167487.2016.12093983>.
- Duckworth, Angela. *Grit: The Power of Passion and Perseverance*. New York: Scribner, 2016.

- DuFour, Richard. "What Is a Professional Learning Community?" *Educational Leadership* 61 (2004): 6–11.
- Education for Seapower: Final Report, December 2018*. Washington, DC: Department of the Navy, 2018.
- Eison, James A. "A New Instrument for Assessing Students' Orientations Towards Grades and Learning." *Psychological Reports* 48, no. 3 (1981): 919–20. <https://doi.org/10.2466/pr0.1981.48.3.91>.
- Eison, James A., Howard R. Pollio, and Ohmer Milton. "Educational and Personal Characteristics of Four Different Types of Learning- and Grade-oriented Students." *Contemporary Educational Psychology* 11, no. 1 (1986), 55. [https://doi.org/10.1016/0361-476X\(86\)90012-3](https://doi.org/10.1016/0361-476X(86)90012-3).
- Ellinger, Emily, and Mark N. Posard. *Imagining the Future of Professional Military Education in the United States: Results from a Virtual Workshop*. Santa Monica, CA: Rand, 2023. <https://doi.org/10.7249/CFA2148-1>.
- Elliott, Stephen N., Thomas R. Kratochwill, Joan Littlefield Cook, and John F. Travers. *Educational Psychology: Effective Teaching, Effective Learning*. Boston, MA: McGraw-Hill, 2000.
- Elliott, Victoria. "Thinking about the Coding Process in Qualitative Data Analysis." *Qualitative Report* 23, no. 11 (2018): 2855. <https://doi.org/10.46743/2160-3715/2018.3560>.
- Enlisted Professional Military Education Policy*, CJCS Instruction 1805.01C. Washington, DC: Chairman of the Joint Chiefs of Staff, 2021.
- Ennis, Robert H. "Critical Thinking Assessment." *Theory into Practice* 32, no. 3(1993): 179–86. <https://doi.org/10.1080/00405849309543594>.
- Examination of the U.S. Air Force's Science, Technology, Engineering, and Mathematics (STEM) Workforce Needs in the Future and Its Strategy to Meet Those Needs*. Washington, DC: National Academies Press, 2010.
- Eyler, Joshua R. *How Humans Learn: The Science and Stories Behind Effective College Teaching*. Morgantown: West Virginia University Press, 2018.
- Felten, Peter. "Principles of Good Practice in SoTL." *Teaching & Learning Inquiry* 1, no. 1 (2013): 121–25. <https://doi.org/10.2979/teachlearninqu.1.1.121>.
- Ferlazzo, Larry. "A Professional Learning Community Is Not a Faculty, Grade Level, or Department Meeting." *Education Week*, 27 April 2021.
- Fiddler, Morris, and Catherine Marienau. "Developing Habits of Reflection for Meaningful Learning." *New Directions for Adult and Continuing Education*, no. 118 (Summer 2008): 76. <https://doi.org/10.1002/ace.297>.
- Flavell, John H. "Metacognition and Cognitive Monitoring: A New Area of Cognitive-developmental Inquiry." *American Psychologist* 34, no. 10 (1979): 908–10. <https://doi.org/10.1037/0003-066X.34.10.906>.
- Fleith, Denise de Souza, and Cristiano Mauro Assis Gomes. "Students' Assess-

- ment of Teaching Practices for Creativity in Graduate Programs.” *Avaliação Psicológica: Interamerican Journal of Psychological Assessment* 18, no. 3 (2019): 306–15. <https://doi.org/10.15689/ap.2019.1803.15579.10>.
- Flom, Jacalyn M., Karen M. Green, and Steven Wallace. “Helping Higher Education Students Succeed: An Examination of Student Attributes and Academic Grade Performance.” *Active Learning in Higher Education* 24, no. 2 (2013): 221–35. <https://doi.org/10.1177/14697874211055689>.
- Frymier, Andrea Bainbridge, Gary M. Shulman, and Marian Houser. “The Development of a Learner Empowerment Measure.” *Communication Education* 45 no. 3 (July 1996): 187. <https://doi.org/10.1080/03634529609379048>.
- Futris, Ted, D. G. Schramm, E. W. Richardson, and T. K. Lee. “The Impact of Organizational Support on the Transfer of Learning to Practice.” *Children and Youth Services Review* 51 (2015): 41–42. <https://doi.org/10.1016/j.childyouth.2015.01.019>.
- Gallagher, Sean R. *Educational Credentials Come of Age: A Survey on the Use and Value of Educational Credentials in Hiring*. Evanston, IL: Center for the Future of Higher Education and Talent Strategy, Northeastern University, 2018.
- Garrett, Richard, Bethany Simunich, Ron Legon, and Eric E. Fredericksen. *CHLOE 7: Tracking Online Learning from Mainstream Acceptance to Universal Adoption: The Changing Landscape of Online Education*. Annapolis, MD: Quality Matters, 2022.
- Glesne, Corrine. *Becoming Qualitative Researchers: An Introduction*, 3d ed. Upper Saddle River, NJ: Pearson, 2016.
- Greene-Shortridge, Tiffany M., Thomas W. Britt, and Carl Andrew Castro. “The Stigma of Mental Health Problems in the Military.” *Military Medicine* 172, no. 2 (2017): 157–61. <https://doi.org/10.7205/MILMED.172.2.157>.
- Greer, Katie, and Sean McCann, “Everything Online Is a Website: Information Format Confusion in Student Citation Behaviors,” *Communications in Information Literacy* 12, no. 2 (November 2018): 150–65.
- Hake, Richard R. “Interactive-Engagement versus Traditional Methods: A Six-Thousand-Student Survey of Mechanics Test Data for Introductory Physics Courses.” *American Journal of Physics* 66, no. 1 (1998): 64–74. <https://doi.org/10.1119/1.18809>.
- Hamilton, Mark. “Prioritizing Active Learning in the Classroom.” *Journal of Military Learning* (October 2019): 3–9.
- Harman, Jane, Eric Edelman, Gen John M. Keane, Thomas G. Mahnken, Mara Rudman, Mariah Sixkiller, Alissa Starzak, and Roger Zakheim. *Report of the Commission on the National Defense Strategy*. Washington, DC: Commission on the National Defense Strategy, 2024.
- Harouni, Houman. “Unprepared Humanities: A Pedagogy (Forced) Online.”

- Journal of Philosophy Education* 55, no. 4–5 (2021): 633–48. <https://doi.org/10.1111/1467-9752.12566>.
- Hartnett, Rodney T., and Warren W. Willingham. “The Criterion Problem: What Measure of Success in Graduate Education?” *Applied Psychological Measurement* 4, no. 3 (1980): 284–85. <https://doi.org/10.1177/014662168000400301>.
- Hasselbladh, Hans, and Karl Yden. “Why Military Organizations Are Cautious about Learning?” *Armed Forces & Society* 46, no. 3 (2020): 475–94. <https://doi.org/10.1177/0095327X19832058>.
- Hayes, Kathryn N., Christine L. Bae, Dawn O'Connor, and Jeffery C. Seitz. “Beyond Funding: How Organizational Resources Support Science Professional Learning.” *American Journal of Education* 126, no. 3 (2020): 389–422. <https://doi.org/10.1086/708266>.
- Hennessey, Megan J. “How Can We Know if Professional Military Education Works?” *War on the Rocks*, 3 August 2020.
- Hennessey, Megan J., ed. *Developing Military Learners' Communication Skills Using the Scholarship of Teaching and Learning*. Maxwell AFB, AL: Air University Press, 2022.
- Hill, Andrew. “Military Innovation and Military Culture.” *Parameters* 45, no. 1 (2015): 86–98. <https://doi.org/10.55540/0031-1723.2809>.
- Hinck, John M. “Designing an Organizational Leadership Development Program: A Case Study Approach.” *Journal of Nonprofit Education and Leadership* 7, no. 4 (2017): 268–86. <https://doi.org/10.18666/JNEL-2017-V7-I4-8081>.
- Hinck, John M., Robert S. Hinck, Steven B. Davis, Derrick J. Iwanenko, Bryan K. Wong, and Kathryn R. Cantu. “Understanding ‘Leadergogy’: The Unique Teaching and Learning Methods in a Leader Development Course.” *Journal of Military Learning* (October 2023): 22–45.
- Hinck, John M., and Steven B. Davis. *Going Virtual: Evolution of the Student Experience Ecosystem (SEE) Model to the SEE Model 2.0 and the Virtual (vSEE) Model*, Eaker Papers no. 4. Maxwell Air Force Base, AL: Air University Press, 2022.
- . “Re-Operationalizing and Measuring Impact of a Leader Development Course.” *International Journal of Teaching and Learning in Higher Education* 32, no. 3 (2020).
- Horning, Alice S. “Critical Reading Skills: An Urgent Challenge.” *Inside Higher Ed*, 16 October 2022.
- Hostetter, Stephanie Teague. “Developing a Quality Assurance Approach for an Online Professional Military Education Institution.” *Journal of Applied Instructional Design* 11, no. 2 (2022). <https://doi.org/10.51869/112/sth>.
- Houser, Marian, and Andrea Bainbridge Frymier. “The Role of Student Characteristics and Teacher Behaviors in Students' Learner Empowerment.”

- Communication Education* 58, no. 1 (January 2009): 35–53. <https://doi.org/10.1080/03634520802237383>.
- Hoyt, Donald P., and Eun-Joo Lee. *Basic Data for the Revised IDEA System*, IDEA Technical Report 12. Manhattan: Kansas State University, 2002.
- Huff, Anne Sigismund. *Designing Research for Publication*. Thousand Oaks, CA: Sage, 2009.
- Huntington, Samuel P. *The Soldier and the State: The Theory and Politics of Civil-Military Relations*. Cambridge, MA: Belknap Press, an imprint of Harvard University Press, 1957.
- Huri, Daman, Noor Haslina, and Mageswary Karpudewan. “Evaluating the Effectiveness of Integrated STEM-Lab Activities in Improving Secondary School Students’ Understanding of Electrolysis.” *Chemistry Education Research and Practice* 20, no. 3 (2019): 495–508. <https://doi.org/10.1039/c9rp00021f>.
- Hutchings, Pat. “Theory: The Elephant in the Scholarship of Teaching and Learning Room.” *International Journal for the Scholarship of Teaching and Learning* 1, no. 1 (2007): 2. <https://doi.org/10.20429/ijSOTL.2007.010102>.
- Iammartino, Ronald, John Bischoff, Christopher Willy, and Paul Shapiro. “Emergence in the U.S. Science, Technology, Engineering, and Mathematics (STEM) Workforce: An Agent-Based Model of Worker Attrition and Group Size in High-Density STEM Organizations.” *Complex & Intelligent Systems* 2, no. 1 (2016): 23–34. <https://doi.org/10.1007/s40747-016-0015-7>.
- Improving the Air Force Scientific Discovery Mission: Leveraging Best Practices in Basic Research Management: A Workshop Report*. Washington, DC: National Academies Press, 2015.
- Jackson, Kimberly, Katherine L. Kidder, Sean Mann, William H. Waggy II, Natasha Lander, and S. Rebecca Zimmerman. *Raising the Flag: Implications of U.S. Military Approaches to General and Flag Officer Development*. Santa Monica, CA: Rand, 2020. <https://doi.org/10.7249/RR4347>.
- Jarvis, Paul, ed. *The Theory and Practice of Teaching*, 2d ed. New York: Routledge, 2006.
- Jeste, Dilip V., Sarah A. Graham, Tanya T. Nguyen, Colin A. Depp, Ellen E. Lee, and Ho-Cheol Kim. “Beyond Artificial Intelligence: Exploring Artificial Wisdom.” *International Psychogeriatrics* 32, no. 8 (2020): 993–1001. <https://doi.org/10.1017/S1041610220000927>.
- Jeste, Dilip V., and Ellen E. Lee. “The Emerging Empirical Science of Wisdom: Definition, Measurement, Neurobiology, Longevity, and Interventions.” *Harvard Review of Psychiatry* 27, no. 3 (May/June 2019): 127–40. <https://doi.org/10.1097/HRP.000000000000205>.

- Johnson, Olenda E. "Creating the Space to Think: The What, Why, and How of Deliberate Reflection for Effective Leadership." *Journal of Character and Leadership Development* 7, no. 1 (Winter 2020): 20–32.
- Johnson, Roy I. "The Persistency of Error in English Composition." *School Review* 25, no. 8 (October 1917): 555–80. <https://doi.org/10.1086/436796>.
- Johnson, Scott D., and Steven R. Aragon. "An Instructional Strategy Framework for Online Learning Environments." *New Directions for Adult and Continuing Education*, no. 100 (Winter 2003): 31–43. <https://doi.org/10.1002/ace.117>.
- Jordan, Jeremy, and Christopher Curtis. "Evaluating the Impacts of Technology Education on Military Maintenance Students." *Journal of Stem Teacher Education* 47, no. 2 (2010).
- Jordan, J. R. "Grief After Suicide: The Evolution of Suicide Postvention." In *Death, Dying, and Bereavement: Contemporary Perspectives, Institutions, and Practices*. Edited by J. M. Stillion and T. Attig. New York: Springer, 2015.
- Kaempf, Sebastian, and Carrie Finn. "Teaching International Relations through the Format of a Massive Open Online Course (MOOC)." *International Studies Perspectives* 22, no. 1 (2021): 1–24. <https://doi.org/10.1093/isp/ekz025>.
- Kebritchi, Mansureh, Angie Lipschuetz, and Lilia Santiago. "Issues and Challenges for Teaching Successful Online Courses in Higher Education: A Literature Review." *Journal of Educational Technology Systems* 46, no. 1 (2017): 4–29. <https://doi.org/10.1177/0047239516661713>.
- Keenlyside, Emily Grace. "Critically Engaging Volunteer Guides: A Study of Adult Learning in an Art Museum." *Canadian Journal for the Study of Adult Education* 31, no. 2 (2019): 8–11.
- Kember, David, Amber Ho, and Celina Hong. "The Importance of Establishing Relevance in Motivating Student Learning." *Active Learning in Higher Education* 9, no. 3 (2008): 249–63. <https://doi.org/10.1177/1469787408095849>.
- Kendeou, Panayiota, Jason L. Braasch, and Ivar Bråten. "Optimizing Conditions for Learning: Situating Refutations in Epistemic Cognition." *Journal of Experimental Education* 84, no. 2 (2015): 245–63. <https://doi.org/10.1080/00220973.2015.1027806>.
- Kent, Thomas. "The 'Remapping' of Professional Writing." *Journal of Business & Technical Communication* 21, no. 1 (2007): 12–14. <https://doi.org/10.1177/1050651906293506>.
- Khachadoorian, Angelle A., Susan L. Steen, and Lauren B. Mackenzie. "Metacognition and the Military Student: Pedagogical Considerations for Teaching Senior Officers in Professional Military Education." *Journal of Military Learning* (April 2020): 6.
- Kilgour, Peter, Daniel Reynaud, Maria Northcote, Catherine McLoughlin, and

- Kevin P. Gosselin. "Threshold Concepts about Online Pedagogy for Novice Online Teachers in Higher Education." *Higher Education Research & Development* 38, no. 7 (2019): 1417–31. <https://doi.org/10.1080/07294360.2018.1450360>.
- Kim, Insu, and Jo Won-kwang. "The Effects of Joint Education Programs on Cohesion among South Korean Military Academy Cadets." *Korean Journal of Defense Analysis* 26, no. 4 (2014): 517–18.
- Kim, Jeonghyun, Lingzi Hong, Sarah Evans, Erin Oyler-Rice, and Irhamni Ali. "Development and Validation of a Data Literacy Assessment Scale." *Proceedings of the Association for Information Science and Technology* 60, no. 1 (October 2023): 620–24. <https://doi.org/10.1002/pr2.827>.
- Knowles, Malcolm. *Andragogy in Action: Applying Modern Principles of Adult Learning*. San Francisco, CA: Jossey-Bass, 1985.
- . *The Modern Practice of Adult Education: From Pedagogy to Andragogy*, 2d ed. New York: Cambridge Books, 1980.
- Kolb, Alice Y., and David A. Kolb. "Experiential Learning Theory: A Dynamic, Holistic Approach to Management Learning, Education and Development." In *The SAGE Handbook of Management Learning, Education and Development*. London: Sage, 2009. <https://doi.org/10.4135/9780857021038.n3>.
- Kolb, David A. *Experiential Learning: Experience as the Source of Learning and Development*. Englewood Cliffs, NJ: Prentice-Hall, 1984.
- Koretz, Daniel M. *Measuring Up: What Educational Testing Really Tells Us*. Cambridge, MA: Harvard University Press, 2009.
- Kremkow, Jennifer M. D., and Erinn H. Finke. "The Effectiveness of a Distance Peer Mentor Training Program with Military Spouses with Children with Autism." *Journal of Autism and Developmental Disorders* 50, no. 4 (2020): 1097. <https://doi.org/10.1007/s10803-019-04334-0>.
- Ku, Kelly Y. L., and Irene T. Ho. "Metacognitive Strategies that Enhance Critical Thinking." *Metacognition and Learning* 5, no. 3 (2010): 264. <https://doi.org/10.1007/s11409-010-9060-6>.
- Kumar, Swapna, Florence Martin, Kiran Budhrani, and Albert Ritzhaupt. "Award-winning Faculty Online Teaching Practices: Elements of Award-winning Courses." *Online Learning Journal* 23, no. 4 (2019): 160–80. <https://doi.org/10.24059/olj.v23i4.2077>.
- Lesar, Timothy S., Laurie Briceland, and Daniel Stein. "Factors Related to Errors in Medication Prescribing." *JAMA* 277, no. 4 (1997): 312–17. <https://doi.org/10.1001/jama.1997.03540280050033>.
- Lewis, Emily. "Best Practices for Improving the Quality of the Online Course Design and Learners Experience." *Journal of Continuing Higher Education* 69, no. 1 (2021): 61–70. <https://doi.org/10.1080/07377363.2020.1776558>.

- Lochmiller, Chad R., and Jessica N. Lester. *An Introduction to Educational Research: Connecting Methods to Practice*. Thousand Oaks, CA: Sage, 2017.
- Loeng, Svein. "Various Ways of Understanding the Concept of Andragogy." *Cogent Education* 5, no. 1 (January 2018): 1–15. <https://doi.org/10.1080/2331186x.2018.1496643>.
- Lukes, Laura A., Sophia Abbot, Dayna Henry, Melissa Wells, Liesl Baum, Kim Case, Edward J. Brantmeier, and Lindsay Wheeler. "Impact of a Regional Community of Practice for Academic Developers Engaged in Institution-level Support for SoTL." *International Journal for Academic Development* 29, no. 3 (2023): 393. <https://doi.org/10.1080/1360144X.2022.2135005>.
- Makopoulou, Kyriaki, Ross D. Neville, Nikos Ntoumanis, and Gary Thomas. "An Investigation into the Effects of Short-Course Professional Development on Teachers' and Teaching Assistants' Self-Efficacy." *Professional Development in Education* 47, no. 5 (2021): 790–91. <https://doi.org/10.1080/19415257.2019.1665572>.
- Manning, Brian, and Kristan J. Wheaton. "Making 'Easy Questions' Easy: The Difficulty of Intelligence Requirements." *International Journal of Intelligence and CounterIntelligence* 26, no. 3 (2013): 597–611. <https://doi.org/10.1080/08850607.2013.780562>.
- Martin, Bella, and Bruce Hanington. *Universal Methods of Design: 100 Ways to Research Complex Problems, Develop Innovative Ideas, and Design Effective Solutions*. Beverly, MA: Rockport Publishers, 2012.
- Mathany, Clarke, Katie M. Clow, and Erin D. Aspenlieder. "Exploring the Role of Scholarship of Teaching and Learning in the Context of the Professional Identities of Faculty, Graduate Students, and Staff in Higher Education," *Canadian Journal for the Scholarship of Teaching and Learning* 8, no. 3 (December 2017): 1–16, <https://doi.org/10.5206/cjSOTL-rcacea.2017.3.10>.
- Maydeu-Olivares, Alberto. "Maximum Likelihood Estimation of Structural Equation Models for Continuous Data: Standard Errors and Goodness of Fit." *Structural Equation Modeling: A Multidisciplinary Journal* 24 (2017): 390. <https://doi.org/10.1080/10705511.2016.1269606>.
- McDowall, Ailie, Carmen Mills, Katie Cawte, and Jodie Miller. "Data Use as the Heart of Data Literacy: An Exploration of Pre-service Teachers' Data Literacy Practices in a Teaching Performance Assessment." *Asia-Pacific Journal of Teacher Education* 49, no. 5 (2021): 487–502. <https://doi.org/10.1080/1359866X.2020.1777529>.
- McKenney, Susan, and Thomas Reeves. *Conducting Educational Design Research*. New York: Routledge, 2012.
- McKinney, Kathleen. "The Scholarship of Teaching and Learning: Past Lessons,

- Current Challenges, and Future Visions.” *To Improve the Academy* 22, no. 1 (June 2004): 8. <https://doi.org/10.1002/j.2334-4822.2004.tb00399.x>.
- McLuhan, Marshall, Quentin Fiore, and Jerome Agel. *The Medium Is the Massage: An Inventory of Effects*. New York: Bantam Books, 1967.
- McMillan, S. L. “Bravo for Brevity: Using Short Paper Assignments in International Relations Classes.” *International Studies Perspectives* 15, no. 1 (2014): 109–20. <https://doi.org/10.1111/insp.12003>.
- Means, Barbara, Yukie Toyama, Robert Murphy, Marianne Bakia, and Karla Jones. *Evaluation of Evidence-Based Practices in Online Learning: A Meta-Analysis and Review of Online Learning Studies*. Washington, DC: Office of Planning, Evaluation, and Policy Development, Department of Education, 2009.
- Merriam, Sharan B., and Elizabeth J. Tisdell. *Qualitative Research: A Guide to Design and Implementation*. San Francisco, CA: John Wiley and Sons, 2009.
- Meyers, Steven A. “Using Transformative Pedagogy When Teaching Online.” *College Teaching* 56, no. 4 (2008): 219–24. <https://doi.org/10.3200/CTCH.56.4.219-224>.
- Mezirow, Jack. *Fostering Critical Thinking in Adulthood: A Guide to Transformative and Emancipatory Learning*. San Francisco, CA: Jossey-Bass, 1991.
- . *Learning as Transformation: Critical Perspectives on a Theory in Progress*. San Francisco, CA: Jossey-Bass, 2000.
- . *Transformative Dimensions of Adult Learning*. San Francisco, CA: Jossey-Bass, 1991.
- Mezirow, Jack, and Edward W. Taylor. *Transformative Learning in Practice: Insights from Community, Workplace and Higher Education*. San Francisco, CA: Jossey-Bass, 2009.
- Miller, Tom, Jonathan Tice, and Tommy Brabson. “Metacognitive Reflection: The Framework for Facilitating Reflective Practice during the Coast Guard Midgrade Officer and Civilian Transition Course.” *Journal of Military Learning* (October 2021).
- Miller-Young, Janice, and Nancy L. Chick, eds. *Becoming a SoTL Scholar*. Elon, NC: Elon University Center for Engaged Learning, 2024. <https://doi.org/10.36284/celelon.oa6>.
- Mitchell, Tom, and Alice M. Cahill. “Cognitive Style and Plebe Turnover at the U.S. Naval Academy.” *Perceptual and Motor Skills* 101, no. 1 (2005): 55–62. <https://doi.org/10.2466/pms.101.1.55-62>.
- Mulready-Stone, Kristin. “A New Form of Accountability in JPME: The Shift to Outcomes-Based Military Education.” *Joint Force Quarterly* 112, no. 1 (2024): 30–38.

- Mundschenk, Nancy A., and Wendy W. Fuchs. "Professional Learning Communities: An Effective Mechanism for the Successful Implementation and Sustainability of Response to Intervention." *Southeastern Regional Association of Teacher Educators Journal* 25, no. 2 (2016): 55–64.
- Nabavi, Razieh Tadayon. "Bandura's Social Learning Theory and Social Cognitive Learning Theory." *Theory of Developmental Psychology* 1, no. 1 (2012).
- Nadelson, Louis S., and Anne L. Seifert. "Integrated STEM Defined: Contexts, Challenges, and the Future." *Journal of Educational Research* 110, no. 3 (2017): 221–23. <https://doi.org/10.1080/00220671.2017.1289775>.
- National Science Foundation: *A Workforce Strategy and Evaluation of Results Could Improve Use of Rotating Scientists, Engineers, and Educators*. Washington, DC: Government Accountability Office, 2018.
- Naval Education Strategy 2023. Washington, DC: Department of the Navy, 2023.
- Neal, Jennifer Watling, and Zachary P. Neal. "Nested or Networked?: Future Directions for Ecological Systems Theory." *Social Development* 22, no. 4 (November 2013): 728. <https://doi.org/10.1111/sode.12018>.
- Newcomer, Jason, Aaron M. Glassman, Maj Aisha D. DaCosta-Paul, and MSgt Jason A. Fowler. "Just Checking the Box: Do Our Airmen Value Their CCAF Degree." *Air & Space Power Journal* 30, no. 1 (2016): 39–42.
- Ng, Davy Tsz Kit, Jac Ka Lok Leung, Kai Wah Samuel Chu, and Maggie Shen Qiao. "AI Literacy: Definition, Teaching, Evaluation and Ethical Issues." *Proceedings of the Annual Meeting of the Association for Information Science and Technology* 58, no. 1 (2021). <https://doi.org/10.1002/pra2.487>.
- Nilson, Linda. *Teaching at Its Best: A Research-based Resource for College Instructors*, 4th ed. San Francisco, CA: Jossey-Bass, 2016.
- Nilson, Linda B., and Ludwika A. Goodson. *Online Teaching at Its Best: Merging Instructional Design with Teaching and Learning Research*. New York: John Wiley & Sons, 2017.
- Novak, Katie, and Catlin R. Tucker. *UDL and Blended Learning: Thriving in Flexible Learning Landscapes*. Oxford, UK: Impress Publishing, 2021.
- Oakley, Barbara, Debra Poole, and MaryAnne Nestor. "Creating a Sticky MOOC." *Online Learning Journal* 20, no. 1 (March 2016): 1–12. <https://doi.org/10.24059/olj.v20i1.731>.
- Officer Professional Military Education Policy, Chairman of the Joint Chiefs of Staff Instruction 1800.01F. Washington, DC: Joint Chiefs of Staff, 2020.
- O'Meara, Kerry Ann. "Ensuring Equity Service Work." *Inside Higher Ed*, 9 May 2018.
- Outcomes-Based Military Education Procedures for Officer Professional Military Education. Washington, DC: Chairman of the Joint Chiefs of Staff, 2022.
- Parenteau, Danic. "Teaching Professional Use of Critical Thinking to Officer-

- Cadets: Reflection on the Intellectual Training of Young Officers at Military Academies.” *Journal of Military Learning* (April 2021).
- Perla, Peter P., and Ed McGrady. “Why Wargaming Works.” *Naval War College Review* 64, no. 3 (2011): 111–30.
- Phillips, John LaForest. “Not All Pain Is Gain: Lessons from Teaching Critical Thinking Online.” *PS: Political Science and Politics* 54, no. 1 (2021): 178. <https://doi.org/10.1017/S1049096520001584>.
- Posner, George J., Kenneth A. Strike, Peter W. Hewson, and William A. Gertzog. “Accommodation of a Scientific Conception: Toward a Theory of Conceptual Change.” *Science Education* 66, no. 2 (1982): 211–27. <https://doi.org/10.1002/sce.3730660207>.
- Potter, Ellie Melissa. “Perceptions of Creativity among Faculty in Higher Education.” Master’s thesis, Grand Valley State University, 2013.
- Preventing Suicide in the U.S. Military: Recommendations from the Suicide Prevention and Response Independent Review Committee*. Washington, DC: Suicide Prevention and Response Independent Review Committee, Department of Defense, 2022.
- Quible, Zane K. “Impact of Error Labeling on Error Elimination in Business Writing.” *Business Communication Quarterly* 69, no. 1 (2006): 8–24. <https://doi.org/10.1177/1080569905285597>.
- Rainville, Kristin N., Cynthia G. Desrochers, and David G. Title, eds. *Faculty Learning Communities: Communities of Practice that Support, Inspire, Engage and Transform Higher Education Classrooms*. Charlotte, NC: IAP, 2014.
- Reigeluth, Charles M., and Yunjo An. “What’s the Difference Between Learning Experience Design and Instructional Design?” *Journal of Applied Instructional Design* 12, no. 3 (January 2023): 227–40. <https://doi.org/10.59668/515.12897>.
- Reis, Harry T., Michael R. Maniaci, Peter A. Caprariello, Paul W. Eastwick, and Eli J. Finkel. “Familiarity Does Indeed Promote Attraction in Live Interaction.” *Personality and Social Psychology* 101, no. 3 (2011): 557–70. <https://doi.org/10.1037/a0022885>.
- Report of a Workshop on Science, Technology, Engineering, and Mathematics (STEM) Workforce Needs for the U.S. Department of Defense and the U.S. Defense Industrial Base*. Washington, DC: National Academies Press, 2011.
- Review of Specialized Degree-Granting Graduate Programs of the Department of Defense in STEM and Management*. Washington, DC: National Academies Press, 2014.
- Rosen, A. M., and L. Kerr. “Wargaming for Learning: How Educational Gaming Supports Student Learning and Perspectives.” *Journal of Political Science Education* 20, no. 2 (2024): 318–35. <https://doi.org/10.1080/15512169.2024.2304769>.

- Rotenberg, Robert. *The Art & Craft of College Teaching*, 2d ed. New York: Routledge, 2016.
- Roth, Wolff-Michael, and Yew-Jin Lee. "Contradictions in Theorizing and Implementing Communities in Education." *Educational Research Review* 1, no. 1 (2006): 27–40. <https://doi.org/10.1016/j.edurev.2006.01.002>.
- Ruel, Erin. "What Is Construct Validity?" In *100 Questions (and Answers) About Survey Research*. Thousand Oaks, CA: Sage, 2019. <https://doi.org/10.4135/9781506348803.n55>.
- Ryan, Thomas G. "The Scholarship of Teaching and Learning within Action Research: Promise and Possibilities." *Inquiry in Education* 4, no. 2 (November 2013).
- Sadler, Philip M., and Gerhard Sonnert. "Understanding Misconceptions: Teaching and Learning in Middle School Physical Science." *American Educator* 40, no. 1 (2016): 26–32.
- Saldaña, Johnny. *The Coding Manual for Qualitative Researchers*. Los Angeles, CA: Sage Publications, 2016.
- Sandwith, Bree L., David H. Cropley, and Lisa J. Chantler. "The Influence of Cognitive Structure and Task Structure on Creativity in a Military Context." *International Journal of Creativity and Problem Solving* 27, no. 2 (2017): 95–112.
- Sanger, Catherine Shea, and Nancy W. Gleason, eds. *Diversity and Inclusion in Global Higher Education: Lessons from Across Asia*. New York: Palgrave Macmillan, 2020.
- Scardamalia, Marlene, and Carl Bereiter. "Computer Support for Knowledge-Building Communities." *Journal of the Learning Sciences* 3, no. 3 (1994): 265–83. https://doi.org/10.1207/s15327809jls0303_3.
- Schatz, Sae, David T. Fautua, Julian Stodd, and Emilie A. Reitz. "The Changing Face of Military Learning." *Journal of Military Learning* (April 2017).
- Schatz, Sae, and J. J. Walcutt. "Modeling What Matters: AI and the Future of Defense Learning." *Journal of Defense Modeling and Simulation: Applications, Methodology, Technology* 19, no. 2 (2022). <https://doi.org/10.1177/15485129221088718>.
- Schraw, Gregory, and Rayne Sperling Dennison. "Assessing Metacognitive Awareness." *Contemporary Educational Psychology* 19, no. 4 (1994): 473–74. <https://doi.org/10.1006/ceps.1994.1033>.
- Sexton, J. Bryan, Eric J. Thomas, and Robert L. Helmreich. "Error, Stress, and Teamwork in Medicine and Aviation: Cross Sectional Surveys." *Journal of Human Performance in Extreme Environments* 6, no. 1 (2001): 6–11. <https://doi.org/10.7771/2327-2937.1019>.
- Shadish, William, Thomas D. Cook, and Donald Thomas Campbell. *Experimental*

- and *Quasi-experimental Designs for Generalized Causal Inference*. Boston, MA: Houghton Mifflin, 2002.
- Sharp, Marie-Louise, Nicola T. Fear, Roberto J. Rona, Simon Wessely, Neil Greenberg, Norman Jones, and Laura Goodwin. "Stigma as a Barrier to Seeking Health Care among Military Personnel with Mental Health Problems." *Epidemiologic Reviews* 37, no. 1 (2015). <https://doi.org/10.1093/epirev/mxu012>.
- Shi, Dexin, Taehun Lee, and Alberto Maydeu-Olivares. "Understanding the Model Size Effect on SEM Fit Indices." *Educational and Psychological Measurement* 79, no. 2 (2019): 310–34. <https://doi.org/10.1177/0013164418783530>.
- Shin, Sungwon, and Jongpil Cheon. "Assuring Student Satisfaction of Online Education: A Search for Core Course Design Elements." *International Journal on E-Learning* 18, no. 2 (2019): 147–64.
- Smith-Osborne, Alexa. "An Intensive Continuing Education Initiative to Train Social Workers for Military Social Work Practice." *Journal of Social Work Education* 51, no. suppl (2015): 89–90. <https://doi.org/10.1080/10437797.2015.1001290>.
- Sonner, Brenda S. "A Is for 'Adjunct': Examining Grade Inflation in Higher Education." *Journal of Education for Business* 76, no. 1 (2000): 5–8. <https://doi.org/10.1080/08832320009599042>.
- Stage, Frances K., and Kathleen Manning, eds. *Research in the College Context: Approaches and Methods*, 2d ed. New York: Routledge, 2016.
- Stake, Robert E. *The Art of Case Study Research*. Thousand Oaks, CA: Sage, 1995.
- State Department: *Additional Actions Needed to Address IT Workforce Challenges*. Washington, DC: Government Accountability Office, 2022.
- Steele, John, Rick Holbeck, and Jean Mandernach. "Defining Effective Online Pedagogy." *Journal of Instructional Research* 8, no. 2 (2019): 5–8.
- Stewart-Wells, A. Gillian, and K. Mallery Keenan. "Assessing Doctoral Students: A Background Comprehensive and Authentic Assessments." *Journal of Continuing Higher Education* 68, no. 2 (2020): 85. <https://doi.org/10.1080/07377363.2020.1743950>.
- Sun, Taiyi. "Forced Experimentation: Teaching Civic Engagement Online amid COVID-19." *PS: Political Science and Politics* 54, no. (2020): 1176–79. <https://doi.org/10.1017/S1049096520001559>.
- Tofade, Toyin S., Jamie Elsner, and Stuart T. Haines. "Best Practice Strategies for Effective Use of Questions as a Teaching Tool." *American Journal of Pharmaceutical Education* 77, no. 7 (2015): 2. <https://doi.org/10.5688/ajpe777155>.
- Truckenbrod, Nathan. *Data Literacy: A Necessary Pre-Condition for Making Data Driven Decisions*. Fort Eustis, VA: Army Training and Doctrine Command, 2023.

- Tzu, Sun. *The Art of War*. New York: Dover Publications, 2002.
- Vincent, Andrea S., Brian P. Decker, and Michael D. Mumford. "Divergent Thinking, Intelligence, and Expertise: A Test of Alternative Models." *Creativity Research Journal* 14, no. 2 (2002): 163–78. https://doi.org/10.1207/S15326934CRJ1402_4.
- Vosniadou, Stella, ed. *International Handbook of Research on Conceptual Change*. London: Routledge, 2013.
- Vygotsky, Lev S. *Thought and Language*. Translated by Alex Kozulin. Cambridge, MA: MIT Press, 1986.
- Walcutt, J. J., and Sae Schatz, eds. *Modernizing Learning: Building the Future Learning Ecosystem*. Washington, DC: Government Publishing Office, 2019.
- Walters, E. M. "Wargaming in Professional Military Education Challenges and Solutions." *Journal of Advanced Military Studies* 12, no. 2 (2021): 81–114. <https://doi.org/10.21140/mcu.20211202003>.
- Warfighting*, Marine Corps Doctrinal Publication 1. Washington, DC: Headquarters Marine Corps, 1997.
- Webb, Andrea S., and Ashley J. Welsh. "Phenomenology as a Methodology for Scholarship of Teaching and Learning Research." *Teaching & Learning Inquiry* 7, no. 1 (2019): 168–81. <https://doi.org/10.20343/teachlearninqu.7.1.11>.
- Winteler, Adolf. "The Academic Department as an Environment for Teaching and Learning." *Higher Education* 10 (1981): 25–35.
- Witty, Paul A., and Roberta La Brant Green. "Composition Errors of College Students." *English Journal* 19, no. 5 (May 1930): 388–93. <https://doi.org/10.2307/803405>.
- Wong, Joseph T., and Bradley S. Hughes. "Leveraging Learning Experience Design: Digital Media Approaches to Influence Motivational Traits that Support Student Learning Behaviors in Undergraduate Online Courses." *Journal of Computing in Higher Education* 35 (2023): 595–632. <https://doi.org/10.1007/s12528-022-09342-1>.
- Wong, Leonard, and Stephen J. Gerras. *Changing Minds in the Army: Why It Is So Difficult and What to Do about It*. Carlisle, PA: U.S. Army War College Press, 2013.
- Yin, Robert K. *Case Study Research: Design and Methods*, 5th ed. Thousand Oaks, CA: Sage, 2014.
- Youssef, Alaa, Samantha Stein, Justin Clapp, and David Magnus. "The Importance of Understanding Language in Large Language Models." *American Journal of Bioethics* 23, no. 10 (2023): 6–7. <https://doi.org/10.1080/15265161.2023.2256614>.

- Zawacki-Richter, Olaf, and Insung Jung, eds. *Handbook of Open, Distance and Digital Education*. Singapore: Springer, 2023.
- Zhang, Xiaodong. "Teaching Citations/Referencing: How Do Chinese College Student Writers Respond?" *Publishing Research Quarterly* 34, no. 4 (2018): 580–94. <https://doi.org/10.1007/s12109-018-9615-y>.
- Zimmerman, Barry J. "Becoming a Self-regulated Learner: An Overview." *Theory into Practice* 41, no. 2 (Spring 2002): 65. https://doi.org/10.1207/s1543042tip4102_2.



INDEX

- Air Force Tuition Assistance (AF-TA), 146–47, 149, 154
- Andragogy, 3, 23–44, 97, 104, 112, 117, 160–62, 166, 185
- andragogical approaches, 84–100, 107, 111, 157, –59, 162, 174, 178, 184–86, 188, 190–210
- artificial intelligence (AI), 25, 190, 192, 196, 203–7, 252, 261, 266, 270, 323–44
 - Open AI, 190–91, 194, 197
 - narrow (ANI), 206, 323–24
 - weak, 323
- assessment, 211–344
- Aviation Weather Cadets, 299–322
- Basic Military Training (BMT), 142–43
- case study research (CSR), 17–18, 164
- certainty of response index (CRI), 300, 304, 306–7, 313, 319–22
- Chief Digital and Artificial Intelligence Office (CDAO), 260–62
- cognitive apprenticeship, 109–11, 116
- cognitive interviews, 59–60
- constructivism, 17, 125, 157, 160–62
- continuum of military learning, 6
- creative learning communities, 5, 15, 19, 25–26, 30, 36, 39, 41, 84–97, 121, 385
- creative problem solving, 47–83
- data literacy assessment, 323–44
- debate, 93–97
- design-based research (DBR), 211–43
- differential item functioning (DIF), 300, 304, 307, 314–15
- divergent thinking, 51–55, 60–62, 68, 70–72, 74, 76–82
- documentation styles
 - American Psychological Association (APA), 102, 108, 117
 - Chicago Manual of Style* (CMS), 102, 108, 117, 119

- Modern Language Association (MLA), 102, 108, 117
- domestic learning framework, 84–86, 88, 90
- École de Guerre*, 84–97
- École Militaire*, 84
- Education Participation Scale (EPS), 136, 138, 147–51, 154
- empathy, 10, 13–14, 24, 26, 36, 39–41, 46, 204, 378
- Enlisted Professional Military Education Policy*, 279
- ethics, 10, 12–13, 24, 37–39, 46, 120–35, 252–54, 261, 324, 326–27, 333–34, 336, 340–43
- experiential learning, 15–16, 18, 21, 24, 29–30, 34, 37, 43, 46, 123, 323–44
- futures seminar, 192–97, 206–7, 330–35
- generative artificial intelligence (AI), 190–210
- generative pretrained transformer (GPT), 190, 194–95, 197, 323–25, 334
- immersive learning environments (ILEs), 25
- immersive learning experience (ILx), 26
- immersive learning simulation (ILS), 25–28
- institutional review board (IRB), 102, 105, 110, 149, 180, 284, 287, 340–41
- instructional techniques, 15–16, 158, 187
- instructor upgrade (IUP) training program, 157–89
- item response theory (IRT), 300, 304, 306–7
- Joint professional military education (JPME), 212, 214, 277–79
- Journaling, 125–35
- language models
 - multimodal large (MLLMs), 191
 - natural language processing (NLP), 339
 - proprietary large (LLMs), 190, 197, 201, 324–25, 333, 338
- learning methodologies, 30, 36, 86, 185
- LOGO-II, 108, 112, 114–15
- machine learning (ML), 252, 261, 325, 329–31, 337
- metacognitive awareness inventory (MAI), 58, 81
- military authoritarianism, 51, 53–55, 60–62, 68, 70, 72, 76–79, 82
- mixed-methods design, 47, 56, 102, 107, 111, 116, 141–42
- mixed reality exercise (MRx), 25–26, 43, 46
- multimodal instruction, 122, 184–85, 188
- Naval Studies Certificate, 244–45, 247, 251–57, 263, 265–66
- Office of Research Compliance, 180
- Officer Professional Military Education Policy* (OPMEP), 278–81, 288
- Open AI, 334
- oral comprehensive exams, 271–98

- outcomes-based military education (OBME), 212–15, 231, 234, 242–43, 271, 278–79, 297
- pinnacle of standards, 33–34, 36
- postevent training, 10, 13, 20, 24, 41–43, 46
- primary developmental education (PDE), 213–14
- professional development, 11, 39, 109, 134, 136–56, 219, 241–42
- program learning outcomes (PLOs), 271–98, 266
- reflection in learning, 123–29
- Scholarship of Teach and Learning (SOTL), 3–5, 101–2, 104–6, 110–11, 115–19, 189, 207, 309
- Service schoolhouses, U.S.
 - Air Force Academy, 4, 39, 299–322
 - Air University, 4–5, 9, 143, 211, 213
 - Air Command and Staff College (ACSC), 11–12, 29–30, 214, 222–23
 - Air War College (AWC), 11–12, 29, 214, 238–39
 - Army War College, 120, 133–23, 125, 128, 192–94, 199, 221, 277–78, 280–96
 - Command and General Staff School (CGSS), 275–77, 280, 295
 - Community College of the Air Force (CCAF), 144
 - Marine Corps University (MCU), 3–4, 48, 247, 258–59
 - Command and Staff College (CSC), 47–48, 50, 54–55, 63, 65–66, 68–71, 73–74, 76, 78, 82
 - Military Academy at West Point, 52
 - Naval Academy, 4, 52, 102, 115, 247, 255, 258
 - Naval Community College, 244–70
 - Naval Postgraduate School, 248, 255, 258
 - Naval War College, 258
- simulation and Wargaming (S&W), 258–59
- Social Cognitive Learning Theory (SCLT), 31–32
- Southern Association of Colleges and Schools Commission on Colleges (SACSCOC), 12–13, 213
- student experience ecosystem (SEE), 30–33
- Survey of Meteorology Concepts (SMC), 299–322
- teaching and learning community of practice (TLCoP), 9–10, 15, 19–20, 23–25, 44–46
- technology acceptance, 329–30, 333
- unimodal instruction, 157–89
- writing error propensity, 101–19



ABOUT THE AUTHORS

DR. MARY L. BARTLETT serves as an associate professor for Air War College at the Leadership and Innovation Institute, Maxwell Air Force Base, AL. She has a PhD in counselor education from Auburn University, and a master of arts degree in counseling from the University of Maryland-College Park. Her principal areas of research are resilience and suicide prevention, and she is an international speaker and published author in those areas. With 27 years of teaching and clinical experience, she was invited to leverage her knowledge in social sciences across all Air Force professional military education. As a subject matter expert in the field of suicidology, Dr. Bartlett works closely with the U.S. Air Force and has contributed to developing policies, strategies, and standard operating procedures while lecturing across all organizational levels of the Air and Space Forces. Prior to her appointment to Air University in 2015, she served for 10 years as a mental health consultant and trainer for leadership across the Department of Defense, advocating and bringing attention to the topic of suicide prevention. Since her appointment to Air University, Dr. Bartlett has held various academic positions including at Air Command and Staff College and the Leadership Development Course.

<https://orcid.org/0009-0000-0226-901X>

MAJOR LOUIS CASCINO hails from Buffalo, NY, where he completed high school before joining the U.S. Air Force. He was commissioned in May 2011, after graduating from the U.S. Air Force Academy with a bachelor's degree in meteorology. He later earned a master's in meteorology from the Naval Postgraduate School in Monterey, CA. During his career, Major Cascino flew global

missions on both the Boeing C-17 Globemaster and the Boeing E-3 Sentry (a.k.a. airborne warning and control system, or AWACS). He also served as a meteorology instructor at the U.S. Air Force Academy. In June 2024, he transitioned from active duty and now serves as a traditional reservist with the 970th Airborne Air Control Squadron at Tinker Air Force Base, OK, while also working as a pilot for United Airlines.

<https://orcid.org/0009-0009-5943-9425>

DR. ANDREW CLAYTON is an assistant professor at Air University within the Leadership and Innovation Institute, Air War College. Dr. Clayton teaches and researches topics related to leadership and experiential learning. He advises and works with several research groups focusing on the human domain in emotional intelligence, empathy, resilience, diversity, and inclusion using immersive technology. Dr. Clayton also provides guidance and research for Air University's Innovation and Technology Research Task Force on current topics related to augmented, virtual, artificial intelligence, and mixed-reality immersive learning environments. He holds a master's in adult education and training, is certified as a master instructor and distance learning instructor, is certified in the Myers-Briggs Type Indicator (MBTI), and is certified as an Air Force innovator. Dr. Clayton also holds a doctorate degree in organizational leadership with an emphasis in higher education and specializes in using immersive technology for leadership and human domain applications.

<https://orcid.org/0009-0001-0276-152X>

EMILIE ALICE CLERET is a Franco-British academic specializing in transformative pedagogy for leaders. She serves as senior advisor to the president of the French Defence University and chief academic officer for English studies in French higher military education, overseeing programs at the *École de Guerre* and the *Centre des hautes études militaires*. She is also the director of debate. With 24 years of experience, she has designed talent development programs in English for military leadership, managed teaching and administrative teams, fostered links between France, the United States, and the United Kingdom (UK), and led seminars on public speaking and transformative learning. Since 2017, she has represented France at NATO's Bureau of International Language Coordination (BILC), managing collaborations with the United States and UK. As a RUSI associate fellow, she has authored articles on transformative learning and given talks on leadership and critical thinking. She has received the *Palmes Académiques* and the *Ordre National du Mérite*.

<https://orcid.org/0009-0007-0918-8715>

JÉRÔME COLLIN is head of the French department in the French War College. He recruits and manages a team of around 15 teachers who teach French as a foreign language to the foreign officers who have been selected to study at the French War College. He also teaches French to the foreign officers in the *Centre des hautes études militaires* and to the defense attachés. He succeeded in the competitive exam to prepare for the Modern Literature Aggregation at the *École Normale Supérieure*. He then graduated from the Université de Paris and holds a master's degree in literature and a master's degree in semiotics. He is preparing for a PhD in semiotics on *l'attente fiduciaire et les structures contractuelles non verbalisées*.

<https://orcid.org/0009-0002-0301-2384>

DR. MARK J. CONVERSINO is professor emeritus and former provost and chief academic officer at Air University. He retired on 31 May 2024, following more than 42 years of service as a career military officer and then as a civilian employee of the Air Force. While at AU, from 2008 to 2015, Dr. Conversino served as the dean of the Air War College, having joined the school's faculty following his retirement from active duty with the Air Force. From 2015 to 2019, he was professor of strategy and deputy commandant of AU's elite School of Advanced Air and Space Studies. In 2019, he assumed the role of AU's provost and chief academic officer. In this position, he served as the university's senior executive authority and primary advisor to the AU commander on all matters related to curriculum, faculty, regional and specialized accreditation, and institutional research and effectiveness. Dr. Conversino holds a bachelor's degree in history from Eastern Kentucky University and a master's and PhD in history from Indiana University, specializing in military history as well as Russian history and Soviet and East European politics. He taught in numerous programs within AU and the DOD, focusing on contemporary Russian military affairs and foreign policy as well as national security strategy. He is the author of the book *Fighting with the Soviets: The failure of Operation Frantic, 1944–1945* (1997), which was later published in Ukraine in 2015. He also published numerous chapters in edited works as well as articles and reviews in academic and professional defense journals. A former member of the board of advisors to the Baltic Defence College in Tartu, Estonia, Dr. Conversino has traveled extensively across the former Soviet Union during the past two decades and in retirement, continuing to lecture and teach in programs across AU and the Department of the Air Force.

DR. CHUCK CUSHMAN is dean of academics at the U.S. Naval Community College (USNCC). Prior to joining USNCC in 2021, Dr. Cushman was at

National Defense University as associate provost and was dean of the College of International Security Affairs. He was a senior fellow at Georgetown University's Government Affairs Institute. He was associate dean and program director at George Washington's Graduate School of Political Management. He was a defense consultant for Headquarters, U.S. Air Force; the Office of the Secretary of Defense; and the Space Commission. In the 105th Congress, he was legislative assistant to David E. Price (D-NC), working on defense and foreign policy. Cushman is a graduate of the U.S. Military Academy, earning a BS in international relations. He served for nine years on active duty as an armor officer. He earned his MA and PhD in American politics at the University of North Carolina in Chapel Hill.

<https://orcid.org/0009-0007-3151-2526>

DR. STEVEN B. DAVIS is an assistant professor of leadership at the Leadership and Innovation Institute, Air War College, Air University. He teaches AU's Leader Development Course, where he also serves as director of faculty development and director of curriculum. He teaches a variety of courses on leadership and history. He completed his BA in sociology at Texas A&M University in College Station, an MA in global history from Universität Bremen, Germany, and he returned to Texas A&M to complete his PhD in history. His publications include leadership education teaching and learning methodologies, leadership course design, faculty development, and twentieth century Central European war and society.

<https://orcid.org/0000-0003-0994-9952>

RUSSELL EVANS is a retired U.S. Navy surface warfare officer. His first opportunity to teach was at the Navy's Aegis Training and Readiness Center, Dahlgren, VA. Since then, he has accumulated 15 years in the classroom split between the Navy and the Marine Corps on active duty, as a contractor, and as a government employee. His first post-secondary teaching opportunity was at Command and Staff College, Marine Corps University, where he was both faculty and curriculum manager and still occasionally adjuncts. He first worked in distance education at Marine Corps University, where he helped build its continuing education program as an academic officer. He is currently the associate dean of naval studies for USNCC. He has studied at the Naval Academy (BS history), George Washington University (MA public policy), and Naval War College (MA national security and strategy).

<https://orcid.org/0009-0009-3659-3745>

MAJOR CHRIS FRANCIS is a senior instructor in the Department of Physics and Meteorology at the U.S. Air Force Academy. He is a 2010 distinguished graduate of Air Force Reserve Officer Training Corps, Detachment 002, and holds advanced degrees in physics and education. He has served with the Air Force Research Laboratory, Air Force Reserve Officer Training Corps, and the Arnold Engineering Development Complex.
<https://orcid.org/0009-0009-1351-7870>

LIEUTENANT COMMANDER PHILIP GARROW, EdD, is a career U.S. Navy surface warfare officer, who completed guided missile cruiser, frigate, littoral combat ship, and destroyer squadron tours. His most significant leadership positions were as operations officer during two afloat deployments to combat zones and as the associate chair of the U.S. Naval Academy's English Department. He holds a BA from Tulane University, LA; MA degrees from Salve Regina University, the U.S. Naval War College, and the University of Maryland-College Park; and a doctorate from the Johns Hopkins University School of Education. His primary research focus is human performance.
<https://orcid.org/0000-0002-9871-7290>

SENIOR MASTER SERGEANT NATHANIEL C. GEORGE is the superintendent of the Department of Physics and Meteorology at the U.S. Air Force Academy. He oversees the management of the Meteorology Lab, audio-visual equipment, and provides technical support for the department. As the superintendent, he plays role in advising leadership, overseeing administrative processes, and ensuring the morale and welfare of personnel within the unit. He enlisted in May 2007 and completed Vehicle Operations Technical School in August 2007. He was first stationed at Royal Air Force Mildenhall, UK, as a vehicle operator and dispatcher before moving to Pope Army Airfield, NC, where he served as NCOIC of the pick-up and delivery section. In 2014, he graduated from the Air Force Weather Technical School and was assigned to the 25th Operational Weather Squadron at Davis-Monthan AFB, AZ, supporting NORTHCOM and SOUTHCOM. SMSgt George deployed multiple times in support of Operations Iraqi Freedom and Enduring Freedom.
<https://orcid.org/0009-0002-7110-4176>

DANIEL V. GOFF, USMC (Ret), is a professor and course director at US-NCC. In previous active-duty roles, he served as the squadron weapons and tactics instructor for Marine Heavy Helicopter Squadron 461 (HMH-461), the operations officer and detachment officer in charge for HMH-464, and the com-

manding officer for HMH-366. He was an Olmsted Scholar (Kyiv, Ukraine) and served as the deputy director of the Stockdale Center for Ethical Leadership (U.S. Naval Academy), and as an analyst with the Strategic Initiatives Group (Commandant's staff). His studies of politics, ethics, and philosophy can be found in publications such as the *Politics in Gotham: The Batman Universe and Political Thought* (2019), *The Politics of Horror* (2020), the *Journal of Catholic Legal Studies* (2021), the *Marine Corps Gazette* (2022–24), *Expeditions with MCUP* (March 2023), and *Batman's Villains and Villainesses: Multidisciplinary Perspectives on Arkham's Souls* (2023).

<https://orcid.org/0009-0002-5673-7119>

WILSON GONZÁLEZ-ESPADA is a professor of physics and science education in the Department of Engineering Sciences, Morehead State University, KY. His academic background includes a bachelor's degree in physics teaching (University of Puerto Rico Río Piedras) and a master's in science education (Interamerican University of Puerto Rico San Germán). González-Espada completed his PhD in science education from the University of Georgia. His teaching experience spans middle, high school, community college, and university. Since 2001, he has taught courses in introductory physics, physical science for general education and preservice teachers, science methods, research methods, and the history of science and mathematics. González-Espada has more than 60 peer-reviewed publications and two books. His scholarly interests include physics education, multicultural STEM education, educational assessment, science and socioscientific communication, and STEM attrition.

<https://orcid.org/0000-0002-9629-1740>

LAYTON GRAVES, EdD, is an educator, logistician, and military officer who currently resides in North Pole, AK. He is an adjunct professor with the Johns Hopkins University School of Education, where he received his doctor of education degree. Additionally, Graves serves as an active duty logistics readiness officer in the U.S. Air Force, with two overseas deployments and duty stations in Alaska; Washington, DC; and New Mexico. Dr. Graves specializes in creating engaging professional development programs for adult learners and military populations. In his free time, he enjoys making ice cream with his wife and son and hiking with their two dogs.

<https://orcid.org/0009-0004-5547-1456>

DR. SCOTT D. HAMM served 26 years in the U.S. Marine Corps, retiring as a sergeant major. During that time, he had the good fortune to serve with many types of units and also to be involved with educating Marines for Marine Corps

University at several different points and positions in his career. Now at the USNCC as the professor of modern naval history, he continues to enjoy educating servicemembers from the Marines, Navy, and Coast Guard.
<https://orcid.org/0009-0002-2412-0173>

KIMBERLY DE LA HARPE is a professor of physics in the Department of Physics and Meteorology at the U.S. Air Force Academy. She received her BS in chemistry from the University of Portland in 2005 and her PhD in physical chemistry from the Ohio State University in 2011, where she studied the excited-state dynamics of DNA with Professor Bern Kohler. Since joining the Air Force Academy faculty in 2011, her research interests have focused on physics education research and the photophysical characterization of organo-metallic compounds. Dr. de La Harpe previously served as the director of the academy's Center for Physics Education Research.
<https://orcid.org/0000-0002-7294-4057>

DR. MEGAN J. HENNESSEY is the provost at Marine Corps University. She is a member of the editorial board for Marine Corps University Press and a founding editor of the journal *International Perspectives in Military Education*. Dr. Hennessey created the Military Scholarship of Teaching and Learning (MSOTL) Forum with Dr. Lauren Mackenzie (Marine Corps University). Previously, Dr. Hennessey held positions at Air University, the U.S. Army War College, the National Geospatial-Intelligence College, the FBI Academy, and the Expeditionary Warfare School at Marine Corps University. Her research interests include the scholarship of teaching and learning within military learning environments and faculty development. She holds a PhD in education from George Mason University, an MA in English from University College London, and a BA in English and religious studies from Hollins University. She earned a direct commission as a public affairs officer in the U.S. Navy.
<https://orcid.org/0000-0002-9310-2860>

COLONEL CHRISTOPHER JOHN HICKEY, PhD, is an associate professor at the U.S. Army War College (USAWC), teaching Military Strategy and Campaigning and Game Theory for Policy and Strategy. In 2025, he is scheduled to teach the Space Force Senior Leader Education program's Joint Global Campaigning Course at Johns Hopkins School of Advanced International Studies (SAIS). Previously an assistant professor at West Point, his work has been published in *Orbis*, the *Journal of Defense Modeling and Simulation*, and *Military Times*. As a functional area 59 strategist and infantry officer, he served in Korea, Iraq, Afghanistan, Haiti, and Europe. He holds advanced degrees from the University

of Maryland, Harvard Kennedy School, and the USAWC. He is a graduate of the NATO Defense College and Johns Hopkins University. He recently participated in projects in support of Ukraine, Sweden and Finland, and NATO, and a research project on AI and planning at USAWC.
<https://orcid.org/0000-0003-0637-1155>

DR. JOHN M. HINCK serves as the associate professor of leadership and deputy director for the Leadership and Innovation Institute, Air War College, Air University. He received his PhD in leadership studies from the University of San Diego, CA, and is a national/international board-certified coach (PCC with ICF), specializing in executive leadership coaching and organizational change. Hinck has more than 40 publications that span leadership, coaching, teaching/learning, faculty development, communication, and the Indo-Pacific region. His books, *Strength and Honor* and *Badges of Honor*, are collections of leadership stories. He is a former Boeing AH-64D Apache Longbow pilot and Army colonel with multiple combat and Joint tours and has 82 months of command time spanning the company, battalion, and task force levels. He serves on the leadership team of the Military and Communication Division for the National Communication Association and on the board of directors for two local non-profits and the International Society for Exploring Teaching and Learning.
<https://orcid.org/0000-0001-8721-405X>

DR. ROBERT S. HINCK is an associate professor of leadership at the Leadership and Innovation Institute, Air War College, Air University. He also serves as the deputy director of AU's quality enhancement plan and deputy director of research for the Leadership and Innovation Institute. Hinck has a PhD in communication studies from Texas A&M University and holds a graduate certificate in China studies from Texas A&M's Bush School of Government and Public Service. He teaches courses on leadership and international communication at AWC, ACSC, and the Ira C. Eaker Center. His research specializes in strategic narratives, US-China relations, and leadership development. He is lead author of two books, the most recent entitled *The Future of Global Competition: Ontological Security Narratives Chinese, Iranian, Russian, and Venezuelan Media*.
<https://orcid.org/0000-0002-8892-9978>

DR. BRADLEY N. KELLY holds a doctorate of education in learning and organizational change from Baylor University, TX, and is a strategic plans and doctrine staff officer assigned to First Air Force (1 AF)–Air Forces Northern command headquarters at Tyndall Air Force Base, FL. Dr. Kelly is an air battle manager with the Air National Guard and has served in various teaching and

instructing roles. Within the U.S. Air Force, Dr. Kelly has served close to 10 years as a lead air battle manager instructor and chief of operational training as well as instructed courses for multiple professional development programs. Additionally, Dr. Kelly served several years before joining the U.S. Air Force as a licensed public school teacher, instructing various social science classes to both middle and high school students, where he specialized in differentiated instructional techniques and the creation of creative learning environments. <https://orcid.org/0000-0003-3345-3279>

DR. SARA KITSCH is an assistant professor of leadership at the Air War College's Leadership and Innovation Institute. She serves as the codirector of faculty development for the Leader Development Course for squadron command and is a member of Air University's quality enhancement plan committee, specializing in the competency of empathy. Dr. Kitsch has a PhD from Texas A&M University in rhetoric and public affairs and holds a graduate certificate in women's and gender studies. She has published multiple academic articles and coauthored a book, with her research interests spanning topics including rhetorical criticism, strategic narratives, empathy, women in politics, and public memory. At Air University, she teaches courses at AWC, ACSC, the Ira C. Eaker Center, as well as several professional development workshops in empathy, conflict management, and communication. <https://orcid.org/0000-0001-9334-2330>

DR. LAUREN MACKENZIE is Marine Corps University's professor of military cross-cultural competence. She earned her MA and PhD in communication from the University of Massachusetts Amherst and has taught intercultural and interpersonal communication courses throughout the DOD for the past 15 years. Dr. Mackenzie's research is devoted to the impact of cultural differences on difficult conversations, and she has written a range of articles and book chapters pertaining to end-of-life communication, relationship repair strategies, and the role of failure in teaching and learning. She is the coeditor with Dr. Kerry Fosher of the 2021 book, *The Rise and Decline of U.S. Military Culture Programs, 2004–20* and the coeditor with Colonel Dana Perkins of the 2022 and 2024 edited volumes, *Women, Peace & Security in Professional Military Education*. Dr. Mackenzie is the 2020 and 2024 civilian recipient of the Marine Corps University Rose Award for Teaching Excellence. <https://orcid.org/0009-0003-4896-9689>

DR. KATHLEEN MOORE is an associate teaching professor of security and risk analysis at the College of Information Sciences and Technology at the

Pennsylvania State University. She obtained her doctorate in information sciences and technology from Penn State. Dr. Moore is the former professor of data science at the Army War College, where she taught data and AI literacy and decision making, strategic personal cyber defense, and strategic futures analysis. She has also taught cyber intelligence at the University of Maryland, and intelligence studies at James Madison and Mercyhurst Universities. Her research includes data and AI literacy and decision making, analysis of online semantic environments during extreme events, domestic violent extremists in online spaces, and digital trust and credibility. Moore also serves as an expert witness in court cases involving the sociotechnical analysis of social media, and a special consult to the Center for Strategic Leadership at the Army War College.

<https://orcid.org/0009-0000-2991-1025>

MATTHEW MYNK is a first-year undergraduate student at the University of Oklahoma currently studying meteorology with a planned minor in mathematics. He is a graduate of the Craft Academy for Excellence in Science and Mathematics, working on the survey of meteorology concepts project as a student there. Mynk has a research interest in atmospheric instability and severe storms, as well as risk management and social awareness of severe weather events. He is currently applying for field research projects conducted by the University of Oklahoma, examining supercell thunderstorms primarily in late spring. He plans to receive at least a master's degree and eventually apply for an operational forecasting/research job for an institution such as the National Weather Service.

<https://orcid.org/0009-0006-7273-454X>

DR. ROB NYLAND is an assistant professor and director of research at the Global College of Professional Military Education at Air University. Dr. Nyland has worked as a learning engineer, researcher, multimedia designer, and faculty member. He has presented at various military and education conferences and published articles about professional military education, learning analytics, online learning, and open educational resources in such publications as the *Journal of Computing in Higher Education*, *International Review of Research in Open and Distributed Learning*, and *TechTrends*.

<https://orcid.org/0000-0003-3264-7039>

LIEUTENANT COLONEL DANIEL O'KEEFE graduated from the U.S. Air Force Academy in physics and mathematics, and then earned his MS in physics from Purdue University and PhD in applied physics from the Air Force

Institute of Technology. He has served as a physicist in the U.S. Air Force since 2010, with assignments at the Air Force Research Lab Weapons Directorate, the Air Force Nuclear Weapons Center, and Department of Physics and Meteorology at the U.S. Air Force Academy. He is currently a program element monitor for ICBMs at the Pentagon. His research interests have focused on physics education, space, nuclear, and hyperspectral, infrared, polarimetric, and event-based imaging.

<https://orcid.org/0000-0003-1398-9906>

CAPTAIN THOMAS ROBERTSON, USN (Ret), currently serves as the inaugural associate dean of emerging programs at the USNCC. He is also a founding director of the U.S. Naval Academy's Forum on Emerging and Irregular Warfare Studies, where he previously taught as a distinguished military professor. A career naval aviator in the electronic attack community, as well as a Navy scuba diver and naval parachutist, Captain Robertson led on the ground and in the air during multiple combat deployments to Iraq, Afghanistan, and the former Yugoslavia. He holds degrees from the U. S. Naval Academy and Georgetown University's Walsh School of Foreign Service.

<https://orcid.org/0009-0006-4373-5797>

COLONEL HEATHER SMIGOWSKI, PhD, is a U.S. Army strategist and serves as the department chair for distance education at the Army War College, Carlisle Barracks, PA. She is also the General Colin Powell Chair of Military and Strategic Studies and course director for the leadership ethics elective offerings. Receiving her commission as an intelligence officer, Colonel Smigowski has served at all military levels. She has served as a civilian intelligence analyst with the Defense Intelligence Agency and a research fellow with the Rand Arroyo Center. Colonel Smigowski received a bachelor of arts from Western Michigan University and holds master's degrees from Clayton College of Natural Health, Regent University, and the U.S. Army War College. In 2015, she received her doctorate in organizational psychology from Capella University.

<https://orcid.org/0009-0000-1383-7262>

LIEUTENANT COLONEL TIMOTHY SPARKS is the director of Command and Staff College Blended Seminar Program, Quantico Region, Marine Corps University. He is a U.S. Marine with more than 20 years of service, including in Iraq and South Korea. His military occupational specialties are 0102 manpower officer and 8802 training and education officer. He has an MA in philosophy from San Diego State University, an MA in higher education from the University of Michigan, and MA in international relations from Webster

University, and an EdD from Johns Hopkins University. His chapter is an edited version of the needs assessment he conducted as part of his doctoral dissertation.

<https://orcid.org/0009-0005-6755-2437>

CARL P. WATTS is a fellow of the Royal Historical Society. He is an associate professor of national security studies at Air University, Global College of Professional Military Education, where he has served as course director across several programs. He has published many articles and book chapters, including “A Nation at Risk?: Education Debates and Policies from Reagan to Trump,” in *Understanding and Teaching Contemporary U.S. History since Reagan* (2022).

<https://orcid.org/0000-0002-7435-0666>

DR. BRETT D. WEIGLE is an associate professor at the U.S. Army War College and is the course director for military strategy and campaigning and an elective for energy and national security. He has taught in the Army War College’s distance and resident education programs for 15 years and is published in the *Journal of Military Ethics and Gender*; *Work & Organization*; an edited volume on strategic landpower from Army War College Press; and a chapter on energy security in *Sustaining America’s Strategic Advantage* (2023). A retired Army logistics officer, he has served overseas in Bahrain, Germany, Turkey, Spain, and twice in South Korea; his U.S. deployments include Somalia and Bosnia, and with NATO in Kosovo and North Macedonia. He teaches design thinking to organizations outside the Army War College, including the NATO Intelligence Fusion Centre, the Army War College of Nigeria, and the U.S. Cyber National Mission Force.

<http://orcid.org/0000-0003-2170-8377>

KRISTAN J. WHEATON is the professor of strategic futures at the Army War College, where he teaches the Futures Seminar. Designed to help U.S. Army senior leaders think more rigorously and effectively under conditions of deep uncertainty, the Futures Seminar is a six-month, project-based program that looks at political, technical, business, and social trends that could impact the Army. He also manages the Futures Lab, which exposes senior Army leaders to cutting-edge commercial technologies. Finally, Wheaton is the author of the *Sources and Methods* blog, several books including *The Warning Solution* and *Wikis and Intelligence Analysis*, and has supervised more than 150 strategic futures projects for a variety of organizations and companies including U.S. national intelligence agencies, Fortune 500 firms, and local small businesses.

<https://orcid.org/0009-0001-4621-8836>

DR. MATTHEW WOESSNER is the dean of faculty and academic programs at the College of International Security Affairs at National Defense University. He previously served as the associate provost for research and assessment at the Army War College. Dr. Woessner has coauthored numerous studies on the impact of ideology in higher education, including the coauthored book, *The Still Divided Academy: How Competing Visions of Power, Politics, and Diversity Complicate the Mission of Higher Education*. His work on ideology has received hundreds of scholarly citations and been cited in both the national and international press to include: *The Guardian*, *The Washington Post*, *The Wall Street Journal*, and *The New York Times*. From 2001 to 2019, he was a member of the faculty at Penn State University in Harrisburg. Woessner served as the 52d chair of the Penn State University Faculty Senate.

<http://orcid.org/0009-0008-5934-6694>

