12-6-2023

Model for Measuring U.S. Army Warfighter's Critical Thinking Skills: A Meta-Analysis

Danita Kornegay Ladson
Walden University

Walden University

College of Health Sciences and Public Policy

This is to certify that the doctoral dissertation by

Danita K. Ladson

has been found to be complete and satisfactory in all respects,
and that any and all revisions required by
the review committee have been made.

Review Committee
Dr. Janine Allwright, Committee Chairperson,
Public Policy and Administration Faculty

Dr. Olivia Yu, Committee Member,
Public Policy and Administration Faculty

Chief Academic Officer and Provost
Sue Subocz, Ph.D.

Walden University
2023

Abstract

Model for Measuring U.S. Army Warfighter's Critical Thinking Skills: A Meta-Analysis

by

Danita K. Ladson

MS, U.S. Army War College, 2016

MA, Central Michigan University, 1997

BA, Wilberforce University, 1986

Dissertation Submitted in Partial Fulfillment

of the Requirements for the Degree of

Doctor of Philosophy

Public Policy and Administration

Walden University

11 February 2024

Abstract

The United States Chairman of the Joint Chiefs of Staff declared critical thinking as a capability imperative for Force XXI. He directed Military Services to develop critical thinkers. There is limited evidence of success in terms of developing critical thinking abilities. Specifically, the U.S. Army lacks a single, generally accepted, codified method for assessing warfighters' critical thinking skills. Using a quantitative approach with a meta-analysis design, this study aimed to examine the extent to which common components of higher-level thinking (remembering, understanding, applying, analyzing, evaluating, and creating) contribute to developing junior and mid-level learners' critical thinking abilities. Bloom's revised taxonomy model served as the theoretical framework for this research study. Random purposeful sampling was employed. Primary studies identified during the literature search were coded using a preset criterion, yielding 13 qualified studies. Subsequent statistical analyses yielded a medium positive effect size of 0.501 (SE = 0.125, 95% CI = 0.256 to 0.747), confirming that the higher-level thinking features contribute to developing critical thinking skills and are generally effective for creating a critical thinking skills assessment strategy. Heterogeneity results indicated that the variance in effect size cannot be confidently attributed to sampling error (Q = 289.931, df = 39, p = 0.001). The study's outcome offers the U.S. Army and military discrete, structured features most efficacious in evaluating the warfighters' critical thinking abilities. This stance promotes positive social change by informing the U.S. Army of an opportunity to advance the armed forces' cognitive readiness to maintain the nation's defense and fulfill national security requirements.

Model for Measuring U.S. Army Warfighter Critical Thinking Skills: A Meta-Analysis

by

Danita K. Ladson

MS, U.S. Army War College, 2016

MA, Central Michigan University, 1997

BA, Wilberforce University, 1986

Dissertation Submitted in Partial Fulfillment

of the Requirements for the Degree of

Doctor of Philosophy

Public Policy and Administration

Walden University

11 February 2024

Dedication

I dedicate my dissertation work to my family, friends, and our military's great men and women. My mother, Harriette Ladson, you were always there to try to help. My Sister Marcellette Greene and Brother Neal Ladson, you helped push me through. My Significant Other, Kevin Hill, patiently waited for me. My Co-workers, you've been my sounding board. The mightiest Men and Women of our military, you sacrifice and work hard to defend this nation.

Acknowledgments

First and Foremost, I acknowledge God our Provider, for with him, all things are possible. This paper and the research behind it would not have been possible without the treasured support of my Committee Chair, Dr. Janine Allwright. Also, thanks to Dr. Olivia Yu for agreeing to serve on my committee and providing the necessary guidance and support at the right time. I sincerely thank Dr. Charles Guthrie for his invaluable tutelage and influence in shaping my data analysis plan and critiquing my results. My gratitude extends to the Faculty of Walden University, specifically the Public Policy and Administration Program Department and Academic Advising, who did not give up on me. I thank Dr. James Mosko, who drew me to the meta-analysis research design. I sincerely appreciate my loved ones, family, friends, and co-workers for their encouragement, patience, and support throughout this lengthy endeavor.

Table of Contents

List of Tables

List of Figures

Chapter 1: Introduction to the Study

This chapter includes an introduction to the topic of this study, a model for measuring U.S. Army warfighters' critical thinking skills. This subject was examined to assist the U.S. Army and military in developing measures and systematic procedures to evaluate the development and practice of critical thinking abilities within the Armed Services. Due to increasingly complex and dynamic operational environments, senior leadership has stressed the necessity for 21st-century competency (Chairman, Joint Chiefs of Staff [CJCS], 2020; Odierno & McHugh, 2015; Straus et al., 2013). Critical thinking, communication, creativity, and innovation are crucial skills and competencies that employees need to possess to thrive in the 21st century (American Management Association [AMA], 2019). Critical thinking skills are among the most demanded competencies during the 21st century for the military and industries. Academic and training institutions purport to teach critical thinking skills. Still, U.S. Army senior leaders continue to appeal for its development among U.S. Army service members because of uncertainty regarding the quality of their critical thinking skills.

Acquisition of critical thinking skills is a key priority area for the U.S. Army as it strives to enhance the effectiveness of its warfighting personnel. Effective critical thinking assessment strategies will better equip military personnel with cognitive tools to adapt to complex situations. This can improve the U.S. Army's ability to deliver and sustain an unmatched scale of ground forces to advance the protection of the nation. Furthermore, the study might lead to social change by helping to address U. S. Army senior leaders' concerns regarding whether service members are acquiring the necessary critical thinking skills to function as valuable professionals and engaged citizens.

This chapter provides context by sharing the background and purpose of the study. The theoretical framework is explained to provide structure and support for the study. I introduce the methodology and describe using a quantitative meta-analysis design to answer the research question. I then address assumptions, limitations and delimitations, significance of the study, and a summary of the chapter. Limitations and delimitations contribute to contextualizing research findings and strengthening the credibility and validity of the research. This study is significant to the U.S. military and related literature.

Background

To confront global security challenges and adversarial threats against the U.S. and its allies, senior Army leadership connects critical thinking skills with warfighters' moral, physical, and tactical capabilities (Austin, 2021; CJCS, 2020). Due to the ever-changing nature of warfighting, warfighters can no longer rely exclusively on tactical and operational skills to carry out their duties (Austin, 2021; CJCS, 2020; Odierno & McHugh, 2015). Driven by the changing character of warfare, the 38th Chairman of the CJCS has levied Military Services to provide education and complement training to produce professionally competent critical thinkers. The U.S. Army depends on the professional military education system to teach warfighters cognitive skills and mental habits. However, U.S. Army senior leaders signaled uncertainty about whether warfighters possess sufficient critical thinking abilities (CJCS, 2020; Odierno & McHugh, 2015; Straus et al., 2013). More importantly, the U.S. Army does not possess a single generally accepted codified method to assess warfighters' critical thinking skills, regardless of the military context in which soldiers learn the craft.

In any situation, U.S. Army warfighters, regardless of rank or profession, must be able to examine a problem, gather and collect available information, analyze facts and assumptions, extrapolate courses of action, and make sound operational judgments. These are attributes of critical thinkers. The extent to which warfighters have honed their critical thinking skills is unknown.

Current and foreseeable security environments can become complicated, irregular, fluid, dynamic, and unsympathetic. Cognitive demands on U.S. military personnel exceed those of the past (Odierno & McHugh, 2015; Straus et al., 2013). As a result, critical thinking abilities have become a priority for the U.S. Armed Forces. Odierno and McHugh (2015) stated that the Army of 2025 will require warfighters with "warfighting capabilities and an expeditionary mindset comprised of confidence, competence, and critical thinking skills" (p. 10). Additionally, the 38th Army Chief of Staff noted that warfighters will be better equipped and able to adapt to unanticipated scenarios dictated by the environment if they have three abilities: competence, confidence, and critical thinking abilities (Odierno & McHugh, 2015).

The CJCS stressed the importance of the professional military education system's role in cultivating leaders to be competent, forward leaning, and critical thinkers. The CJCS (2020) further indicated that "deterring, fighting, and winning" (p. 2) against adversaries cannot be done without investing in cognitive competencies. The U.S. Army Training and Doctrine Command [TRADOC] Regulation 25-36, which establishes doctrine for the U.S. Army, underlined the value of critical and creative thinking as an enabler for Force XXI, referred to as U.S. Army's operational forces at all echelons. According to Straus et al. (2013), the broad spectrum of operations carried out by warfighters underscores the U.S. Army's need to transform the PMES to promote adaptability and critical thinking skills. TRADOC (2014) cautioned that leadership development must center on higher-order cognitive skills that "cultivate soldiers' ability to dissect information quickly; rapidly assess situations accurately; develop options; and easily transition as the operation changes" (p. 6). Biddle (2016) suggested that the U.S. Army focuses on improving critical thinking, communication, and understanding senior civilian leaders' worldviews and behavior. Biddle considered these attributes essential to help the U.S. Army deal with the challenges of predicting and preparing for the future and addressing problems. Ayers (2016) indicated that the U.S. Army does not know the level of critical thinking talent across the workforce. The U.S. Army lacks measures to determine whether higher education develops critical thinking talent in its leaders (Ayers, 2016). Ayers suggested that the U.S. Army employ the professional military education system to strengthen soldiers' critical thinking skills.

Problem Statement

The U.S. Army depends on the professional military education system to teach warfighters cognitive skills and mental habits of the mind. The repeated requests from Department of Defense authorities to develop the critical thinking abilities of warfighters indicate a need for more confidence in their possession of these abilities (CJCS, 2020; Odierno & McHugh, 2015; Straus et al., 2013). The U.S. Army does not possess a single generally accepted or codified method to assess warfighters' critical thinking skills, regardless of the military context in which they learn the craft. In this context, I attempted to assist the U.S. Army in overcoming its constraints involving assessing critical thinking skills capabilities and potentially extending throughout the general U.S. military population.

This study is relevant, current, and significant to the discipline. While critical thinking is not new, emphasis on developing the skill has gained significance in private and public sectors and academia. Critical thinking is one of the 21st century skills needed to operate in society. Students who aspire to participate in the modern era, which is marked by complex cultures and a globalized economy, must be able to collaborate with others to solve problems, exercise critical thought, communicate clearly, and accept globalization (Care et al., 2018). Although critical thinking skills are emphasized as a 21st century human attribute, clear indicators of levels of competency involving these skills are lacking. These indicators are central to designing assessment frameworks.

Purpose of the Study

This quantitative meta-analysis study involved determining the extent to which conventional features of higher-level thinking—remembering, understanding, applying, analyzing, evaluating, and creating—significantly contribute to developing junior and mid-level learners' critical thinking skills in postsecondary education. Assuming this study demonstrates that these higher-level thinking features contribute to developing critical thinking skills, the U.S. Army professional military education system is set to receive a framework to measure and validate critical thinking skills throughout the

Armed Forces. The findings of this study will deliver value to the Department of the Army [DA], professional military education faculty, and TRADOC. The research will help warfighters improve their learning ability and benefit from critical thinking training and education. Additionally, this research may enhance and broaden future literature to further investigate and address conceptual processes associated with critical thinking. Testing hypotheses, collecting data, and evaluating and interpreting results depend on establishing independent and dependent variables. The independent variables in this study are: remembering, understanding, applying, analyzing, evaluating, and creating.

Research Question

In this study, I used the following research question:

RQ: To what extent can standard features of higher-level thinking (remembering, understanding, applying, analyzing, evaluating, and creating) contribute to developing junior and mid-level learners' critical thinking skills in postsecondary education?

$H_0$1: Standard features of higher-level thinking (remembering, understanding, applying, analyzing, evaluating, and creating) do not contribute to developing critical thinking skills in post-secondary education.

H_a1: Standard features of higher-level thinking (remembering, understanding, applying, analyzing, evaluating, and creating) contribute to the development of critical thinking skills in postsecondary education.

Theoretical Framework

Bloom's revised taxonomy model served as the framework for the study. Bloom's taxonomy is a strategy for educators, philosophers, and psychologists. The taxonomy levels and forms of learning are ordered into six progressive stages: remembering, understanding, applying, analyzing, evaluating, and creating (Facione, 1990). To become skilled in critical thinking, each step must be executed and mastered. Stage one is *remembering*, which involves recognizing or recollecting newly learned content. The remembering stage includes recalling a wide range of information, from specific facts to theories. Stage two is *understanding*, or the capacity to construct meaning from recently learned material by deciphering and translating material or foreseeing results or effects. During stage three, *applying* individuals apply learned material in new and concrete situations. Applying involves using benchmarks, strategies, techniques, ideas, standards, laws, and theories.

The most complicated mental processes are necessary when an individual reaches the following three phases: analyzing, evaluation, and creating. The fourth stage, *analyzing*, adds complexity to the previous stages. Analyzing is the capacity to dissect materials into parts to comprehend organizational structure. This capability involves mental exercises such as differentiating, organizing, and attributing and the aptitude to distinguish between components or groups. Stage five is *evaluation*, which refers to using standards and criteria to determine the material's importance and conduct a thorough and rational data assessment. *Creating* and organizing components into a new pattern or structure is stage six, the last stage in Bloom's revised taxonomy. Creating requires reassembling parts in a new form or product or synthesizing elements into something new (Anderson et al., 2001). Creating is the most complex stage in Bloom's revised taxonomy.

Because it provides an inventory of qualities and characteristics for teaching and assessing critical thinking, Bloom's taxonomy applies to this study. Bloom's revised model emphasizes problem-solving, applications of principles, analytical skills, and creativity (Facione, 1990).

Nature of the Study

This quantitative study involved synthesizing empirical studies with critical thinking skills assessments. I identified conventional features of critical thinking for use in developing an assessment model for the Army to measure warfighters' acquisition of critical thinking skills. I conducted a systematic review using a meta-analysis design. A systematic review is a compilation of systematic approaches designed to convert study results to common metrics and statistically analyze relationships between variables and effects (Glass, 1976). The review helped to determine whether adequate information exists to proceed with a meta-analysis. The system review supported that enough data existed. Therefore, a meta-analysis ensued.

Researchers use meta-analysis to examine performance tests, compare interventions, and assess program effectiveness. Meta-analysis addresses the direction and scale of effects across studies (Wilson, 2010). The general steps of a meta-analysis include: formulation of the problem, establishing criteria, searching for and selecting studies, assessing risk, extracting effect sizes, coding features, synthesizing effect sizes, analyzing data, and interpreting and reporting outcomes (Basu, 2014). I applied these general steps to collect the data, analyze, and formulate findings for this study.

Possible Types and Sources of Data

This research encompassed a comprehensive literature search to identify essential empirical studies relevant to the central research question. Keywords in the review include “critical thinking assessment,” “critical thinking measures,” “teaching critical thinking,” and their variations. The review integrated the cognitive skills signified in Bloom’s revised taxonomy. The electronic databases cited below were some sources used to identify relevant studies. ProQuest Dissertations & Theses Global, Educational Resources Information Center (ERIC), Cochrane Database of Systematic Reviews, PsycINFO, Academia Social Sciences Index, Military and Government Collection, Education Research Complete, Social Science Direct, and Google Web and Scholar. These databases held a selection of peer-reviewed studies, maximizing the sources required to complete this meta-analysis.

A study must fulfill predetermined criteria to be included in a meta-analysis. The following criteria were applied in this study. The selected studies must be accessible, available publicly, or archived; address the issue of critical thinking development, improvement, practice, assessment, and application; involve instruction-based interventions for developing critical thinking compared with another intervention or no intervention; include at least two independent variables, pre-and post-test experimental or quasi-experimental design; possess sufficient statistical information to extract effect size; be published within ten years of the literature review and written in the English language; and have sample sizes with ten or more participants (Hedges and Olkin, 1985). I conducted a thorough quality review of the studies selected using established critical evaluation guidelines and quality checklists. The detailed quality assessment helped investigate heterogeneity and inform decisions on the suitability of meta-analysis. Following Lipsey and Wilson's (2001) suggestions, I developed a coding technique including statistical and theoretical data to transform the characteristics of the included studies in this meta-analysis into variables.

Proposed Analytical Strategies

Creating a pathway to develop a model for the U.S. Army to measure and validate critical thinking skills among its warfighters required examining experiments and interventions to measure junior and mid-level learners' achievement in learning these skills. A systematic analysis followed by a meta-analysis was beneficial. The meta-analysis consisted of a five-stage process involving collecting suitable studies, coding study features, calculating effect sizes, transforming study outcomes to standard metrics to compare results, examining relationships between study characteristics and outcomes, and assessing bias.

The meta-analysis started with critically appraising information from selected studies to confirm they met internal validity criteria. The next step involved coding each study by characteristics such as study source, demographic environment, study method, instructional features, and measurement features. I also considered the possibility of skewed data. Skewed data are detected when the standard deviation is significant compared to the mean size, especially when the data has a maximum and minimum range value (Cochrane, 2017). When faced with skewed data, a sensitivity analysis is performed. An analysis was not required because I did not detect any skewed data.

Another important calculation was the weighted average of the intervention effects in each selected study. An assessment of bias was also part of this meta-analysis. IBM SPSS Version 28 was used for data analysis, including evaluation of mean effect size, subgroup analysis, publication bias assessment, and heterogeneity test.

Definitions of Terms

In this study, I used the following terms, which are defined here:

Cochrane Handbook for Systematic Reviews of Intervention: The Cochrane Handbook is a comprehensive guide that outlines the process of preparing and maintaining systematic reviews on the effects of interventions.

Critical Thinking: Critical thinking is thinking activities that include active and skillful conceptualizing, applying, analyzing, synthesizing, and evaluating information (National Council of Excellence in Critical Thinking (Scriven & Paul 1987). Critical thinking involves deliberately evaluating issues and subjects, understanding logical questioning and reasoning processes, and proficiency in using such approaches (Glaser, 1972). Some prominent features of critical thinking are defining assumptions, focusing on uncertainties, analyzing discussions, asking and answering questions, and evaluating the reliability of sources (Anderson et al. (2001).

Junior and Mid-Level Learners: Individuals who are studying in a capacity between high school and post-secondary education.

Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA): A guide developed by Cochrane authors to standardize the process and documentation for conducting a systematic review.

Professional Military Education System: A five-level professional training, development, and schooling of military personnel. It includes a broad range of initiatives, training, and other programs to assist military personnel in progressing throughout their careers and being ready for higher levels of responsibility.

Assumptions

According to Creswell (2014), assumptions are factors that cannot be verified but are acknowledged as true. I assumed that a systematic review yielded sufficient information for a meta-analysis. I assumed deviations from the Cochrane Handbook for Systematic Reviews (2023) established meta-analysis protocols led to inaccurate conclusions regarding the associations between the variables. I assumed conscious or unconscious biases might influence how data are perceived during the study's data collection and analysis phases. I strictly followed predetermined protocols while collecting data and performing analysis for this study. The protocol enabled me to remain objective and reduce or eliminate assumptions.

Scope and Delimitations

The U.S. Army does not possess a single generally accepted and codified method to assess warfighters' critical thinking skills (Ayers, 2016; Biddles, 2016). This study involved validating standard features of higher-level thinking that contribute to the development of critical thinking skills. A set of structured and isolated standard features of critical thinking skills will provide a framework for U.S. Army professional military education system facilitators to design an effective model to assess warfighters' critical thinking skills. Also, through this research, I hope to inform the general military and educational professionals about findings regarding the plausibility of a holistic validated model to measure the development and practice of critical thinking skills.

Limitations

Inherent in a quantitative meta-analysis are limitations and constraints that might affect study findings. The first limitation was that several studies needed to report the necessary statistics to include in the meta-analysis. The second limitation depended on the included studies' data quality to make valid inferences. The third limitation was the need for more research on critical thinking measures and evaluation techniques in published literature. The fourth limitation that had to be considered was the need for more data, which could have presented various challenges. For example, it reduces the statistical power and representativeness of samples. A fifth limitation involved the possibility of publication bias. Applying quality inclusion and exclusion criteria, a valid statistical methodology, and interpreting results in context with available data allowed for limitations to be overcome.

Delimitations

Delimitations are qualifiers that define the capacity and limits of a study. The participants' education levels had to be at the college/university level equivalent to the stakeholders of this study, U.S. Army junior and midlevel commissioned officers between the ranks of Second Lieutenant (O-1) and Lieutenant Colonel (O-5). This population represented the U.S. Army's future senior leaders. As their levels of responsibility increase, their ability to reason, analyze, and interpret for good decision-making becomes more important. This demographic will be able to develop new habits of the mind through education, training, and practice. O-6 Colonels and above tend to resist changing old habits or reprogramming. Thus, I explored critical thinking measurement tools to support professional military education courses designed to educate junior and midlevel U.S. Army officers.

Another delimitation included narrowing the study to a quantitative meta-analysis design, limiting what could be a more extensive conceptual understanding of the problem in other military services and academic and social science fields. Due to the study's limited scope, time constraints to complete the study, and generalizing the findings proved to be a complicated process.

Significance of the Study

War today, and conceivably in the foreseeable future, demands a change in how U.S. soldiers operate. Regardless of occupation specialty, warfighters perform tasks and make decisions when information, rules, and guidance may be limited. Realizing that critical thinking skills are necessary for soldiers to operate in rapidly changing operational environments and deal with novel situations, the U.S. Army has mandated change in its PMES and training to equip soldiers with these skills. In 2006, the CJCS identified critical thinking as one of the six desired leadership attributes for Joint Force 2020. The U.S. Army implemented a policy to include critical thinking instruction in course curricula throughout its PMES.

Even though the U.S. Army reformed its educational curriculum and training to include teaching critical thinking skills, objectively measuring soldiers' strengths and weaknesses in developing these skills is challenging. A clear understanding of standard features that contribute to developing critical thinking skills might lead to establishing an effective critical thinking assessment strategy to better equip the Armed Forces with cognitive tools to adapt to complex situations. Furthermore, while there is literature on teaching and developing critical thinking, there is little evidence of a universally accepted gold standard assessment technique or best practice.

I aimed to build on broad literature. Abrami et al. (2008) expressed that researchers in the education and psychology fields have created a variety of critical thinking skills methods that differ in format, characteristics, scope, and application.

If the U.S. Army cannot effectively measure soldiers' critical thinking capabilities, this could result in leaders' inability to identify problems from multiple perspectives, make rational decisions, and possibly achieve and maintain a competitive advantage against adversaries.

Summary

This study involved synthesizing empirical studies with critical thinking skills assessments to identify key critical thinking features that will help the U.S. Army develop a model to measure and validate these skills among its warfighters. Chapter 1 included the study's background, rationale, and conceptual framework. I also addressed assumptions and, known limitations and delimitations. The study is significant to senior military leaders, scholars, and social science and education practitioners.

Chapter 2 includes a literature review regarding the general concept of critical thinking, its importance, and how it is learned and assessed in the U.S. Army. Chapter 3 provides information about the nature of quantitative research and strategies and procedures for sampling, participant selection, data collection, and data analysis and interpretation. Chapter 4 includes an analysis of data, results, findings, patterns, relationships, and themes. Chapter 5 includes interpretations of findings, conclusions and recommendations for action and further study.

Chapter 2: Literature Review

Increasing U.S. Army warfighters' critical thinking capability is a prevailing narrative of the CJCS and other senior military officials. U.S. Army senior leaders hold critical thinking skills as an indispensable core capability that warfighters need to respond to adversarial threats and complex global security challenges. The potential security climate might cause the U.S. to face a spectrum of unforeseeable challenges. Challenges include but are not limited to great power competition, threats of aggressive and destructive operations, disruption of alliances and partnerships, cyberattacks, and humanitarian crises (Joint Operating Environment, 2016). To overcome anticipated challenges, political and military objectives must be flexible, dynamic, and multifaceted, skills achieved mainly through applying critical thinking skills. As a result, cognitive abilities, particularly critical thinking, are in high demand for warfighters (CJCS, 2020(Joint Operating Environment, 2016). Critical thinking capabilities provide advantages that warfighters might not possess without adequate instruction and assessment (Facione, 1990; Paul & Elder, 2008).

Despite efforts to embed critical thinking into the U.S. Army professional military education curriculum, senior Army leadership does not know whether soldiers are developing sufficient critical thinking skills to apply them in the operating environment. At the time of this study, the U.S. Army has yet to have a unified, generally acceptable, and scalable codified tool for evaluating critical thinking skills among warfighters. I address theories relevant to the study and methods for measuring critical thinking capabilities among the U.S. Army's commissioned officer population.

Literature Search Strategy

In this study, I used the following search terms: *critical thinking skills and dispositions*, *assessment approaches or models for critical thinking skills*, *critical thinking in the U.S. Army*, *measuring critical thinking skills*, *measuring soft skills*, *evaluating critical thinking skills or soft skills*, *Bloom's taxonomy*, *student learning outcomes*, and *assessing critical thinking skills in the military*.

Most of the studies selected for review were published between 2008 and 2022. I used the following databases: ABI/INFORM Complete, Academic Search Complete, EBSCOHost, ERIC, PsycINFO, ProQuest Central, ProQuest Dissertations and Theses, Research Gate, Social Science Index, SAGE Journals, Science Direct, and Thoreau. In addition, I investigated general literature on formative and summative critical thinking skills assessment interventions using Google Scholar.

Theoretical Foundation

Bloom's revised taxonomy model, used for designing cognitive learning strategies and assessments, served as a framework to help answer the research question. I used Bloom's revised taxonomy thinking levels, cognitive verbs, and knowledge nouns to characterize assessment practices and patterns in selected critical thinking development studies. Bloom's revised taxonomy involves classifying knowledge and cognitive levels and forms of learning and assessment posed in two components: The knowledge component comprises four hierarchical categories: factual, conceptual, procedural, and metacognitive (Anderson et al., 2001). The cognitive component consists of: *remembering, understanding, applying, analyzing, evaluating, and creating* (Anderson et al., 2001). Bloom’s classification system, divided into knowledge and cognitive components, provides learners with a range of thinking levels to demonstrate the depth of their critical thinking skills (Anderson et al., 2001).

Bloom’s revised taxonomy model can be used to classify the level of knowledge and thinking skills that warfighters need in professional development, training, and delivering combat power in support of national security strategies. For example, an Intermediate Level Education instructor who wants to incorporate critical thinking skills in coursework can use higher-order skills such as application and analysis. These skills and dispositions are associated with an upper-level course curriculum equivalent to senior Captains’ and Majors’ experiences and skills. Conversely, lower-level courses such as the Basic Officer Leaders Course for Lieutenants and junior Captains are akin to skills, knowledge, and comprehension. Critical thinking skills are demonstrated in upper-level courses by warfighters who can classify, compare, and contrast information to reach decisions (Anderson et al., 2001). When warfighters effectively think and reason through problems, they demonstrate cognitive skills such as combining, creating, designing, developing, evaluating, justifying, and measuring. Exercises that stimulate knowledge, understanding, and application of thinking abilities in lower-level classes can establish whether soldiers can recite facts and organize information to answer problems by applying basic principles (Anderson et al., 2001). Learning objectives for lower-level courses are frequently served by the phrases: recall, choose, search, define, demonstrate, explain, build, develop, and use. Scholars and educators can use verbs to build conversation questions that address phases of cognitive thought processing and classify students' responses into taxonomy levels (Halawia, 2007).

Bloom's revised taxonomy model is based on the concept that learners master lower-level categories of the pyramid before progressing to higher-level categories (Anderson et al., 2001; Fastiggi, n.d.; Halawia, 2007). Learners gain knowledge, skills, and a better grasp of learning content as they progress through the taxonomy (Fastiggi, n.d.; Halawia, 2007). Verbs can be used to create learning objectives, questions, and tasks and depict what learners should be able to execute with the content they are learning (Fastiggi, n.d.).

Bloom's revised taxonomy model is one of the most used and enduring instruments to guide successful teaching practice and assessment throughout the education and psychology sectors. However, some educational and psychology professionals challenged aspects of the model. The best-known critics of the model are Case (2013), Forehand (2005), and Wineburg & Schneider (2009). The educators argued that Critical thinking should not be bound to a specific classification of mental operation but a set of characteristics or qualities of thinking as demonstrated in any intellectual activity (Case, 2013; Forehand, 2005). Another criticism is that the taxonomy constitutes a succession, not an authentic integration seen in real-life situations. Wineburg and Schneider (2009) claimed that situating the phrase knowledge at the base of the taxonomy pyramid devalued knowledge in learning.

Despite the validity of these concerns, some counterarguments undermined Case, Forehand, Winesburg, and Schneider's claims. The taxonomy includes a framework of core concepts to be considered when developing curriculum, instruction, and assessment (Munzenmaier & Rubin, 2013). By converting standards into a common language, the revised taxonomy is intended to assist educators in terms of comparing what they aim to achieve regarding learning outcomes. Furthermore, Bloom's revised taxonomy model is structured to balance a full range of skills to accommodate varying instructional outcomes and learners from diverse backgrounds (Soozandehfar & Adeli, 2016).

Wineburg and Schnieder's (2009) contention over the positioning of the phrase knowledge in Bloom's revised taxonomy neglected to capture the capacity and scope of the classification as the basis for future actions of the mind (Fastiggi, n.d.). Bloom viewed knowledge as foundational, and situating it at the lower part of the pyramid was his method of focusing on its significance as the building block for achieving higher-level thinking (Bokhove & Campbell, 2020). In the revised taxonomy, Anderson and Krathwohl (2001) re-emphasized the foundational importance of knowledge by adding it as a separate dimension to the cognitive process dimension, with four types of knowledge included: factual, conceptual, procedural, and metacognitive. As a result, knowledge is no longer one level in the new model but rather a dimension that encompasses all six levels and changes in either direction or beyond (Anderson & Krathwohl, 2001; Bokhove & Campbell, 2020). Bloom's revised taxonomy is more than flexible enough to integrate new observations, allowing the potential to expand on previous models.

Despite the criticisms of Bloom's revised taxonomy model, the framework remains applicable for this study because it provides a generally acceptable inventory of qualities and characteristics for teaching and assessing critical thinking (Facione,1990;

Lau, 2018). Bloom's revised taxonomy aims not to explain how the mind acquires new information but to propose a method for teaching and assessing thinking skills. Educators, psychologists, and academic professionals consistently viewed Bloom's revised taxonomy as one of the primary frameworks contributing to educational program improvement in the 21st century (Anderson, 2001; Fastiggi, n.d.; Halawia, 2007; Lau, 2018). Its longevity supports the view of many educators, philosophers, and psychologists that it is still accepted, practical, and helpful (Facione, 1990; Lau, 2018). Bloom's revised taxonomy provides a structure to build a breadth and depth of educational programs and curricula that accomplish learning objectives.

Further, educational professionals find the taxonomy a straightforward, flexible, and comprehensive framework to teach, learn, and measure cognitive skills (Facione, 1990; Radmehr & Drake, 2019). The development of measures to assess levels of critical thinking capabilities within the U.S. Army formations requires understanding the constructs important to cognitive function and readiness. To that end, through layers of process skills, Bloom's revised taxonomy offers a framework that might help the U.S. Army's PME facilitators develop a model to measure and assess the quality of warfighters' learned critical thinking skills. The core cognitive skills described in this study are readily grasped by educators and used for instruction and evaluation.

Critical Thinking Roots

Reflecting on the origin and contribution of critical thinking helps acquire a broader, more concrete interpretation of the fundamentals of critical thinking, perspectives on what has been practiced, and what has been recognized as viable and ineffective ways of applying and assessing critical thinking. A review of origin might also enhance perspective on slight gaps, differences, or similarities to build upon intellectual road maps of the critical thinking skills framers: Socrates, Aquinas, Descartes, Bacon, Machiavelli, Marx, Darwin, Freud, Sumner, Dewey, and Bloom (Florence, 2014; Hitchcock, 2018; Murawski, 2014; & Paul et al., 1997).

Critical thinking can be traced back to Socrates's teachings, which set the foundation for critical thinking through the line of questioning paradigm. Socrates viewed that knowledge develops from questioning rather than instruction (Cooper & Hutchinson, 1997). Obtaining facts, assessing logic and conclusions, presenting fundamental concepts, and explaining inferences have all been critical to Socrates. These behaviors are shown in the various contemporary definitions of critical thinking. To expose individuals' irrational reasoning or unreliable information, Socrates asked questions (Kim & Mejia, 2019). In his style of questioning, Socrates emphasized the need for thinking to permit transparency and logical consistency. Today, educational institutions, schools, and other learning venues still use Socrates's questioning concept. Greek philosophers Plato and Aristotle (the founder of the study of logic) followed Socrates.

Philosopher Thomas Aquinas personified systematic thinking in his work. He influenced society's understanding of the potential effects of reasoning and the need to improve and cross-examine reasoning (Egan, 2005). Aquinas systematically articulated, considered, and responded to them as an essential stage in their development to ensure that his thoughts withstood critical thinking scrutiny. Aquinas's work also taught that thinking critically does not mean rejecting established beliefs, only those that lack a reasonable basis.

René Descartes, the Father of Modern Philosophy, pioneered a new way of thinking in the 17th century, often known as the Age of Enlightenment. Descartes rejected the philosophical views of Aristotle and Aquinas, arguing that the mind has to be specifically and systematically trained to help it think (Florence, 2014). Descartes developed a method of critical thought based on the principle of systematic doubt, meaning that all aspects of thinking should be questioned, doubted, and tested. Moreover, to ensure that individuals accept what is true, they must consciously renounce all the strong yet doubtful convictions they gained through experience and education. Descartes underpinned dissecting problems into parts, deducing inferences, and performing a deliberate, systematic synthesis considering all things (Serrat, 2017).

During the 15th and 16th centuries, the application of critical thinking gained more influence among scholars in Europe. For example, Francis Bacon was concerned with how people mishandled their minds to pursue knowledge (Weeks, 2019). Bacon underscored the risks of leaving the mind to its natural inclinations. He pointed out that utilizing one's devices produces poor thought patterns (called idols) that prompt false or deceptive thinking. Bacon characterized four main errors in mental processing: The term *idols* of the tribe refers to ways in which individuals tend to trick themselves; *idols of the marketplace* refers to ways in which individuals misuse words; *idols of the theater* denotes a tendency to become trapped in conventional thought patterns; and *idols of the* *schools* indicates thinking problems that are caused by illogical rules and subpar instructions (Weeks, 2019).

According to Engel et al. (2016), Bacon stressed the importance of studying the world empirically. Another critical factor about Bacon is that because he was credited as the first to genuinely understand the power of sound inductive reasoning to generate insights, peers and scholars crowned him the father of empiricism (Muntersbjorn, 2003). Bacon abandoned conventional, inflexible ways of classifying knowledge to favor a new, from-the-ground-up idea that relied on experiments to verify or refute theories. Contemporary critical thinking scholars view his book as one of the first primers of critical thinking.

Niccolò Machiavelli led to the entry of modern critical political thought as he assessed politics during the Italian Renaissance. He rejected the notion that the government operated as individuals in positions of authority or great influence (Onion et al., 2018). Instead, he analyzed how government worked and created a framework for political thought that revealed politicians' real agendas and the inconsistencies and irregularities of the harsh, ruthless world of politics at the time.

In the 18th and 19th centuries, Karl Marx, Charles Darwin, and Sigmund Freud extended the concept of critical thought to humanity, society, and culture. Karl Marx concentrated his efforts on addressing capitalism's political and economic issues and the socialist revolution's problems (The Research Group of Socialism and Democracy, 2014). He applied critical thought to critique capitalist class society. A key point of Marx's critique of the political system was that it remained a form of exploitative social structure. Marx's theory held that a capitalist system inherently contains roots of destruction. He maintained that as ownership and capital accumulate, the working class's exploitative conditions deteriorate over time, prompting them to battle against the middle class. Marx's writings contribute to the current analysis of capitalism.

Critical thinking enabled Charles Darwin to effect a distinct change in biology methodology. According to Mayr (2009), Darwin founded evolution biology and, in the process, introduced historicity into science. Darwin considered collecting evidence that did not agree with his previous assumptions, referring to this principle as a golden rule. Critical thinking is exemplified in Darwin's theory of evolution and his notion that the species most adaptable to change, rather than the strongest, will survive.

Sigmund Freud displayed critical thinking in and through his works on the human mind. In applying critical thinking, Freud discovered psychoanalysis (McLeod, 2017). Psychoanalysis is a treatment for those experiencing psychiatric issues in which they are questioned about their emotions to determine the reason for their condition. Freud also developed the Free Association Therapeutic technique, which includes a picture slideshow for patients to express what happens in their minds. Freud extended critical thought by establishing the *concepts of ego, super-ego, and id*. The concept represents the three distinct elements of the human psyche. According to Freud, the *id* is the hidden, inaccessible aspect of our personalities that governs a person's primal desires. The *ego*, the reality principle, makes realistic efforts to appease the fundamental urges. The *super-ego* strives for perfection; it consists of each person's ego ideals and opposes the *id*.

Freud's study on various topics, including sex, religion, women, and society, exemplifies how he used critical thought.

William Graham Sumner was one of the principal researchers investigating the construct of critical thinking. In 1906, Sumner sparked debate in the sociology discipline when he elevated the idea of critical thinking through the context of Darwin’s theory of evolution (Elder & Cosgrove, n.d.). Sumner raised the argument that critical thinking requires cultivation in an environment that will foster its growth. According to Summer, as individuals grow in their thinking they learn to examine evidence, reject prejudice, and perceive everything as subject to change and scrutiny (Elder & Cosgrove, n.d.). In 1906, Sumner published Folkways, which highlighted the risks of social teaching in schools. He argued that school education produces men and women who are molded by a single routine as if manipulated in a lathe unless guided by the highest knowledge and good sense. The mainstream views are rife with generalizations, half-truths, and fallacies. Sumner bolstered critical thinking throughout life and education; he maintained that practical thinking relies on mental habit and power.

At the dawn of the 20th century, John Dewey popularized critical thinking by igniting the modern-day critical thinking movement (Prawat, 2000). According to Dewey, learning is the main objective of schooling, and learning is an outcome of thinking. Dewey introduced reflective thought and inquiry as fundamentals of critical thinking (Rodgers, 2002). Dewey described reflective thinking as functioning, continuous, and cautious thought of an instinctive or perceived form of information as a result of the motives for its development and the farthest implications wherein it culminates. The academic philosopher's definition is featured in present-day narratives concerning the dynamic nature of critical thinking, reason, and inference. Dewey's work also contributed to the theory of *metacognition*, or thinking about thinking, which emerged in the 1970s and opened the way for critical thinking to become a widely acknowledged educational feature. Dewey's work dramatically transformed evaluation techniques and concepts.

Drawing on the wisdom of early Greek, European, and English scholars, Benjamin Bloom, who, through his research and work with students, strongly influenced the field of education. His studies centered on the idea that the home and the classroom can help people reach their full potential, which inspired experimentation and change in the educational system (University of Chicago Chronicle, 1999). Bloom and a group of American Psychological Association colleagues embarked on exchanges that eventually resulted in developing a taxonomy of educational goals, a classification method now referred to as Bloom's Revised Taxonomy Model. The taxonomy depicts significant components in the cognitive domain and elevates a technique for aligning educational goals, curricula, and assessments (Bellis, 2020).

Critical Thinking Skills and Their Importance to the U.S. Army

Complexity, instability, irregularity, and competitiveness are signals that U.S. warfighters must adjust as they execute future military operations and carry out obligations to protect national interests. The ongoing conversations and experiences acquired from the hybrid conflicts in Iraq and Afghanistan underpin the U.S. Army's foreseeable operational environment. Hybrid war is characterized as a series or simultaneous use of non-conventional modes such as irregular warfare, insurgency, terrorism, and proxy wars (Roberts & Lawson, 2019). A typical school of thought holds that the future is a modified version of the present and that one must first understand past and present conflict to comprehend future conflict. As such, the span of contemporary conflicts and key trends indicates what future military interventions might emerge. These factors shape why critical thinking matters and support soldiers' mental flexibility, interoperability in joint operations, the outmatching of the adversary, and the perspective of global competition.

Critical thinking supports soldiers' mental flexibility. John Dewey alluded to mental or cognitive flexibility in his views on education (Prawat, 2000). Dewey implied that to grow and gain experience, individuals must consistently change and readjust to the environment. The dynamic and unpredictable security environment in which soldiers are expected to operate contributes to the idea that exercising mental flexibility is central to encountering future adversaries. Mental flexibility is the ability to shift perspectives and actions when exposed to new or unexpected events and information (Cojocar, 2011). Cognitively flexible warfighters can better comprehend and adapt to changes in the operational environment, forecast, assess future occurrences, and act quickly in unforeseen scenarios. Warfighters adept in critical thinking and mental flexibility can swiftly reorganize while adjusting their responses to shifting operational conditions. Critical thinking supports the advancement of integrating and operating in a joint environment. The nature of future warfare and constrained resources would further motivate U.S. military services and coalition partners to collaborate on a wide range of operations (Joint Operating Environment, 2016). An essential aspect of functioning in a joint force environment is the interchange between the service components and coalition partners, access to service and resources, and enhanced responsiveness. Military scholars characterize the joint force environment as cooperative arrangements among coalition partners or nations, integrating functions, capabilities, and resources to achieve strategic goals, objectives, and interoperability (Reiter, 2012). Soldiers who think broadly and across domains can gain perspective and a shared understanding of other service components and partner nations' cultures, values, and attributes, ultimately leading to consensus around goals. Warfighters with critical thinking abilities are more likely to be flexible, intuitive, open-minded, and agile in a joint environment (Scott, 2016).

Military leaders and strategists suspect that future warfare will compel U.S. forces to contend against adversaries with extensive experience, progressive forces, and comparable or superior technologies (DA, 2019; Becker & DeFoor, 2018; Scott, 2016). The force that outsmarts the other will likely gain the advantage. Armed Forces with critical thinking abilities are better able to acquire, interpret, analyze, and understand massive volumes of data and foresee likely outcomes than their adversaries. Fusing critical thinking skills with the mastery of warfighting capability and superior technology can create the outcomes necessary to exploit adversaries' systems and vulnerabilities.

Cultural Differences and the Development of Critical Thinking

Critical thinking is founded on two competing perspectives on learning. The assumption that humans learn through the passive transfer of information is one viewpoint. Passive transfer occurs when a novice learns something from an expert; they then become knowledgeable (Advance Consulting for Education, 2013). The other idea is that humans consciously and independently construct their knowledge based on professional expertise and experience. Humans objectively create knowledge. This view is known as constructivism, a core belief of English-speaking cultures (Abrami et al., 2008). In cultures that believe learning is generated through passive transfer, the education system is fact-based and noncritical (Advance Consulting for Education, 2013). Students participate in rote memory, descriptive, and narrative activities that involve little to no critical component.

In cultures that believe in constructivism, the instructional process is critical, whereas there is a constant exchange of information between the instructor and student (Abrami et al., 2008). Problem-solving, critical thought, discovery, dialogue, and argument are examples of constructivism that place knowledge under scrutiny and evaluation. Students who learn through passive transfer need help adapting to constructivism and critical thought. Therefore, students require an explicit understanding of the concept, demonstration, and exercises or activities that enable them to gradually improve their critical thinking skills. Bloom's Taxonomy is the framework for forming constructivism and critical thinking. The taxonomy's six distinct learning process stages help bridge the gap between passive transmission and constructivism. As students progress through the levels, the required thinking becomes more complex.

U.S. Army State of Critical Thinking Capability

There is a shared acknowledgment among the U.S. military senior leadership, both past and present, that critical thinking is an important skill needed to operate in the current and future operational environment (JCS, 2020; Davitch & Folker, 2017; DA, 2019). The U.S. Army senior leadership has identified critical thinking skills as a capability gap. The notion that the U.S. Army leadership is unaware of the critical thinking abilities potential inside its formations further complicates the situation. Examining the extent to which common features of high-level thinking contribute to developing critical thinking skills might assist the U.S. Army Professional Military Educators with encouraging measurement approaches to validate warfighters' critical thinking capacities. The ability to measure critical thinking skills might provide confidence that warfighters would be able to recognize and react appropriately to dynamic alterations in the character and conduct of warfare, as well as to ill-structured national security threats.

For the past 20 years, the U.S. Army has been immersed in warfighting that did not resemble the typical conventional war that forces had trained for and were accustomed to fighting. Conventional wars are generally open confrontations between two or more state actors with well-defined forces using familiar guerilla fighting strategies and doctrinal thinking (West, 2014). During the Iraq and Afghanistan campaigns, warfighters engaged in irregular warfare with multiple unidentifiable adversaries and clandestine terrorist organizations (Berry, 2017). In addition to fighting against an ill-defined enemy, the U.S. Army’s mission expanded significantly to building partnerships and nation-building, eventually taking precedence over destroying them. Therefore, it is widely challenging the U.S. Army’s technical and tactical capacity. The 20th CJCS, General Milley, also echoed this point, asserting that the character of warfare is changing, and future conflicts likely involve a blend of conventional and irregular warfare labeled as “hybrid warfare” or “operations in the gray zone” (Smith, 2019, p.3).

While the U.S. Army forces mastered warfighting and advanced weaponry, they needed to enhance other areas of warfare response, including recognizing and adapting to changes in the operational environment; exploiting adversaries’ vulnerabilities; course-correcting; and thoughtful consideration of objectives and the definition of success (Hooker & Collins, 2015; Smith, 2019; West, 2014).

Lessons from the Iraq and Afghanistan conflicts taught the U.S. Army that the character of contemporary conflict is complicated, unusual, and unstable in scale, timing, salience, and politics (Smith, 2019). The lessons also drew attention to cognitive skill gaps that contributed to challenges encountered throughout the campaigns. For example, strategic military officers’ perception that modern technologies would change how wars are fought clouded their evaluation of the operational environment. It hampered the refinement of war plans and strategies (McMaster, 2013).

This widely held viewpoint resulted in unforeseen political consequences that took time to resolve. Second, military leaders underestimated the will and determination of adversaries (Berry, 2017; White, 2014). The perception that the U.S. military had won past conflicts because the adversary was less capable and motivated produced an overconfident attitude (Berry, 2017). Key decision-makers were deceived by this perception, which had yet to be tried or studied with insurgency operations. Third, the U.S. military forces’ inability to rapidly adjust to the insurgents’ tactics and conditions on the ground reduced the effectiveness of operations, which led to increased costs in lives and operational success.

Next, the coalition strategy in Iraq neglected to adequately address the war's political and humanistic aspects (White, 2014). In a 2014 Veteran's Day address, McMaster, the Deputy Commanding General of the U.S. Army Training and Doctrine Command, indicated that many challenges faced in the protracted conflicts in Afghanistan and Iraq were predetermined by fundamentally erroneous thinking about future warfare (McMaster, 2013). Defense experts and Senior Fellows pointed out that many of the challenges encountered by the U.S. military were primarily rooted in faulty assumptions about its capabilities (Berry, 2017; McMaster, 2013; White, 2014).

Notably, after implementing the new Counterinsurgency (COIN) policy, U.S. forces and coalition allies significantly enhanced combat operations. The new strategy and the demands of complex and fluid combat operations necessitated a shift in thinking and cognitive abilities. The strategy demanded well-rounded soldiers capable of adapting to situations, gathering and analyzing data, making timely decisions, and engaging in developing and reshaping the operational environment, all stimulated by critical thinking. In addition to adopting the COIN Strategy, the US Army shifted to the Mission Command doctrine. The doctrine fosters the decentralization of command and control to encourage freedom in decision-making, rapid speed of action, and effort to take effective action within specified limits (Kalimuddin, 2017). Military officers under mission command should adapt and seize opportunities as they arise and conditions change. The capacity for critical thought strengthens mission command because to execute it; Army commanders must have a broad perspective, shared knowledge, and awareness of actions occurring across the operations. They need to operate in a way consistent with the commander's intentions.

The emphasis on developing critical thinking aptitudes is not new for the U.S. Army. The CJCS (2009) required the armed forces to train forces in critical thinking, as the CJCS expected that cognitive abilities would become increasingly important. The Chairman accorded continuous learning, cross-domain cooperation, and critical and creative thinking top priority in the 2019 Vision and Guidance for Professional Military Education and Talent Management. Contrarily, the training approach used by the U.S. Army focuses on rote learning or muscle memory, which is incompatible with critical thinking. The premise behind this training approach is that skill is learned through repetition and eventually becomes second nature (Harman, 2012). The existing and anticipated warfighting environment does not afford the U.S. Army the ability to train in this manner. Rote learning is relevant for an environment that requires little effort to comprehend concepts, connect new information, adapt to changing conditions, and solve complex problems.

The Iraq and Afghanistan conflicts revealed that the true nature of warfighters' responsibilities and their quality in applying military power depends on a required standard of thinking. The state of critical thinking capability in the U.S. Army unearthed that unless the U.S. Army adjusts and adequately addresses the vulnerabilities among soldiers' cognitive skills through some form of measurement, it risks creating a void in assuming the obligations required for future conflicts.

PME and Developing Critical Thinking Skills

Dempsey (2012) challenged the Joint Professional Military Education (PME) to develop warfighters' habits of mind, expressly critical thinking. To appreciate the total value of mission command, General Dempsey pressed military joint education program facilitators to ingrain in warfighters' cognitive capacity. The General expressed that to execute mission command, Army leaders must be able to perceive and expressly communicate, take specific actions, embrace judicious risk, and build trust within the forces. A look into the U.S. Army Professional Education system's practices in teaching critical thinking skills provides a greater understanding of variables that might impact the nature and scope of this study.

PME is an array of continuing education and programs within the U.S. Army intended to develop service members as they progress through their careers and to prepare them for duties of increasing scope and complexity (Kaurin, 2017). It is essential to look closely at the distinction between training and education because they are used interchangeably in the military. Training is the process of inculcating a specific skill or form of action in an individual. Training applies rote learning to perform a particular task or use an object and is generally paired with practical, hands-on experience (Masadeh, 2012). On the other hand, by imparting knowledge, education creates a long-term behavioral transformation in the individual. Individuals gain the ability to critically connect with their surroundings, solve issues, and make decisions due to their education. Hence, education helps warfighters learn how to think and solve problems, while training teaches them what to do. The PME system is the primary program the U.S. Army requires to develop warfighters' habits of the mind (Caine, n.d.). PME supports officers and enlisted personnel, and its structure, content, and purpose vary across rank, organization, function, tasks, and responsibilities (Kaurin, 2017).

PME Learning Models For Developing Critical Thinking

The PME system is the primary establishment to develop critical thinking skills at various levels of warfighters' careers. While senior leaders depend on the PME to develop habits of mind and critical thinking skills, education seems to be abbreviated and limited. The U.S. Army PME system comprises several instructional and learning platforms and approaches to teaching critical thinking. Traditionally, warfighter training focused on muscle memory or repetition mechanics to ensure that soldiers make the right choices as directed (Harman, 2012). This method lacked emphasis on acquiring thinking skills and instead encouraged knowledge in a straightforward, sequential environment and systematic decision-making. Additionally, it did not prepare soldiers for solving novel, multifaceted problems and challenges (DA, 2018a). This methodology outgrew its relevance for training and preparing warfighters to function in a complex, ambiguous, and disorderly environment. Thus, it provoked the TRADOC and Doctrine Command to adopt meaningful learning and active learning theoretical concepts.

Meaningful learning contrasts with rote learning. Educators recognize meaningful learning as the purposeful construction of meaning by connecting concepts and knowledge to previously acquired knowledge and then applying that knowledge to learn more (Allrich, 2017). Like meaningful learning, active learning is a form of learning in which students engage directly or experientially in the learning process (Bonwell &

Eison, 1991). TRADOC's adoption of these concepts follows the U.S. Army philosophy, which urges expanding educational and training programs to develop the warfighters' cognitive processes and establish systems to assess students' learning and institutional performance (DA, 2018a). Meaningful and active learning aligns with Bloom's revised taxonomy because active learning involves engaging higher-order thinking skills such as analysis, synthesis, and evaluation (Mayer, 2002). These higher-order tasks are described in Bloom's taxonomy as the upper levels of this taxonomy.

In consonance with the meaningful learning and active learning concepts, PME instructors cultivate warfighters' critical thinking skills through Applied Critical Thinking (ACT), Groupthink Mitigation (GTM), Experiential Learning Model (ELM), Problem-Solving, and Decision-Making (DA, 2018a). ACT is the process of applying critical thinking skills and disposition with a focus on curiosity, connection, creativity, and communication. It fosters individuals to challenge their thinking. The U.S. Army PME instructors leverage ACT tools in courses to introduce curiosity, inferences, and reasoning through Socratic questioning, brainstorming, and posing complex problems to analyze. However, educators who do not comprehend the fundamentals of critical thinking and do not comprehend how to develop quality, objective questions are probably inefficient at employing the ACT.

GTM is applying tools to foster divergent thinking during problem-solving by incorporating individuals' views into the group before converging on alternative solutions. GTM aims to eliminate or restrict "groupthink" when individuals cannot separate their points of view from the remaining group members and often arrive at conclusions not shared by the other group members. TRADOC finds GTM valuable in developing social relationships and mitigating groupthink through small group exercises, techniques, and feedback.

The Experiential Learning Model (ELM) is the process of learning through experience, reflective observation, and active experimentation (Kolb, 1984). Using this model in the PME environment, learners acquire critical thinking experience through collaboration between the instructors or subject matter experts and the learners. Dialogue aims to stimulate thought and knowledge construction to achieve critical thinking skills. Also, critical thinking experiences emerge from individual and group projects, practical exercises, and group discussions, casting new perspectives on additional issues that need to be analyzed (Pool-Funai & Hansen, 2016).

Problem-solving and decision-making tasks are two more learning modalities employed in the PME to enhance critical thinking abilities in warfighters. Problem-solving involves understanding the situation, predicting, and finding a solution (Headquarters, Department of the Army (HQDA), 2011). By promoting joint problem-solving events, U.S. Army PME instructors use problem-solving techniques to involve learners in reasoning and learning the significance and meaning of what they read. The U.S. Army educational processes incorporate decision-making approaches to enable learners to question their decision-making processes and conclusions.

This section offered a quick overview of the literature on the concepts, approaches, and activities used throughout the U.S. Army Professional Military Education system to promote warfighters' critical thinking skills.

Three Levels of PME Education

The PME provides education at three distinct levels: tactical, operational, and strategic. The PME levels of education are structured to systematically build Officers' expertise and knowledge at the appropriate grade from Commissioning through General/Flag Officer while elevating cognitive growth at every level (DA, 2018c).

Tactical Level

Generally, professional military education for an Officer starts at the tactical level with the Basic Officer Leader Course (BOLC), a two-phased training program intended to produce commissioned officers in the U.S. Army. Prospective officers complete Phase I (BOLC A) either as a cadet (United States Military Academy or Reserve Officers' Training Corps) or as an officer candidate (United States Army, Officer Candidate School) before progressing to BOLC II as Second Lieutenants. Officers receive a branch assignment after commissioning. BOLC II is the first training an officer completes. It is a rigorous, branch-immaterial course that tests junior officers' mental and physical suitability, with most training conducted hands-on in a tactical or field environment (HQDA, 2019). The training primarily focuses on warrior tasks. Pleban et al. (2006) pointed out that soft skills are taught but are mainly abbreviated and limited to small group instruction, counseling, after-action reviews, and risk management. PME cadre evaluates officers' performance in various leadership positions under different conditions and situations (Caine, n.d.). The officers in training also participate in several peer reviews and self-assessments. The skills learned in BOLC are reinforced at the junior officers' first unit assignment, where on-the-job training and exercises build upon initial training. Officers need to effectively know or develop the tactical experience required to perform the core duties so that the ability to switch from a focus on muscle memory to more sophisticated problem-solving will become easier.

Operational Level

Mid-grade officers Majors (O3) and Lieutenant Colonels (O4) complete the Intermediate Level Education (ILE) and the Advanced Operations and Warfighting Course (AOWC), which is operational level training and the third stage of PME for commissioned officers. Officers who attend ILE are typically at the 10-to-13-year point of their career (HQDA, 2019). PME develops mid-grade officers' critical thinking skills at this level by presenting opportunities to think objectively about various topics, analyze and synthesize information, and propose solutions to complicated problems. In addition to preparing officers for battalion and brigade command and senior staff positions, ILE educates and develops officers for joint, interagency, and multinational operations. Critical and creative thinking are important to an officer's success at this level (Caine, n.d.).

Strategic Level

At the Strategic level, Senior officers (Lieutenant Colonel (O5) and Colonel, (O6)) reach the final stage of PME by attending one of the Senior Service Colleges (SSC): Army War College, Air War College, Naval War College, or The National Defense University (HQDA, 2019). In place of SSC, some officers pursue fellowship studies at leading universities such as Harvard, Georgetown, and the NATO Defense College, and interdepartmental courses such as the Advanced Operational Studies program, the Defense Systems Management College, or the Center for Strategic and International Studies (Caine, n.d.). Each SSC instruction stresses strategic thinking, strategic planning, policy, national security decision-making, operational art, and joint and combined military operations. SSC aims to prepare senior officers to formulate and execute military strategies based on joint and service-specific operations. Critical thinking enables these attributes. Investing in critical thinkers can help create leaders who better understand the environment, make connections, integrate operations, predict change, and react quickly to unexpected events (Wolter et al., 2013).

Methods to Teach Critical Thinking

With an end goal of providing the U.S. Army with a roadmap to an effective critical thinking assessment strategy to better equip warfighters with cognitive tools, it is important to obtain as much information about the targeted skill's concept, practices, and standards as possible to build on this study. This section explores expert methods of teaching critical thinking and formats that repeatedly appear in the literature. However, before identifying teaching approaches, it is necessary to examine whether applying critical thinking is skills-based or context-bound. A closer look at these concepts indicates that there is ambiguity and overlap in how academics view critical thinking should be taught, which must be considered when designing assessment measures.

Ennis (2003), Facione (1990), and Halpern (1998) described skills-based as generic traits such as reasoning, judging, making assumptions, and drawing inferences that can be learned apart from any knowledge domains and then transferred or applied across any field or undertaking. Most school curriculums and assessments conceptualize critical thinking as a generic skill. In contrast to Ennis, Facione, and Halpern's view that critical thinking is skills-based, other scholars hold different views. Kuncel (2011) and McPeck (1981) assigned critical thinking to context-bound, noting that information about a subject is necessary for critical thinking. These authors further pointed out that critical thinking differs widely between disciplines. There remain varying perspectives on categorizing critical thinking as a set of generic traits that apply across subject domains or context-bound used within the context of a specific domain (Al-Ghadouni, 2021; Ennis, 1989; Southworth, 2022). However, modern-day philosophers and educators regard critical thinking as a generic trait (Al-Ghadouni, 2021; Liu, Frankel, & Roohr, 2014). There is no commonly agreed view of critical thinking categorization as a collection of common characteristics or context-bound. However, both must be considered in theoretically designing an algorithm to ensure the measurement is accurate and accommodates various teaching formats and approaches.

Ennis's (1989) critical thinking typology identifies four learning approaches commonly used in instructional interventions to develop student's critical thinking skills: general, infusion, immersion, and mixed. In the general approach, CT skills and dispositions are taught independently with no specific subject matter content. While popular, this approach has only improved critical thinking in undergraduate programs if used with other strategies (Eldridge, 2010; Facione, 1990; Moore, 2004). The infusion approach combines in-depth teaching of the topic and deliberates critical thinking principles (Willingham, 2008). This approach identifies and teaches several critical thinking skill types through the learning material. When studying a subject, students are primarily taught how to think, what processes to use, and how best to reach logical conclusions. The immersion approach varies from the infusion approach in that students are encouraged to practice critical thinking through in-depth subject-matter instruction without explicitly referencing critical thinking principles. In effect, students practice and apply their critical thinking skills but will not receive explicit guidance on the principles. The mixed approach combines the general approach with either the infusion or the immersion approach (Ennis, 1989; Willingham, 2008). Students take part in a unique thread of courses that combine in-depth subject matter training that teaches them how to think critically through foundational critical thinking lessons.

Several factors interact to determine the effectiveness of instructional outcomes in teaching critical thinking and learners' acquisition and practice of the skill. A review of the four learning approaches—general, infusion, immersion, and mixed—helps to understand better the types of instructional approaches that might be found in the collection of individual studies proposed for this meta-analysis.

Assessment Formats to Evaluate Critical Thinking

According to literature reviews on the U.S. military and critical thinking abilities, several instructional and evaluation approaches can help learners develop these abilities. A view of the instructional formats' characteristics and the differentiation between them was helpful in fully comprehending the components needed to model a critical thinking skills assessment process.

Classroom Sessions

Classroom Sessions involve continuous facilitation, monitoring discussions, and sharing experiences among instructors and students. Throughout the classroom sessions, instructors will assess students' critical thinking skills through written assignments, papers, and responses to questions that promote discussion and sharing experiences.

Online Discussion Forums

Similar to classroom sessions, instructors use online discussion forums to encourage collaborative sessions by moderating the discussion forums to stimulate critical thinking. Instead of relying upon reading materials to obtain information, students use open-source data to validate online discussion posts (Arend, 2009). Online discussion forums are a reasonable strategy for developing students' understanding by participating in dialogue throughout the discussion thread (Shana, 2009). Online discussion forums can also help students organize and devise ideas before posting comments. Through interacting with other understudies, students explore the learning process and debate discussion topics using analysis and supporting evidence.

CLS

CLS facilitates groups working together to solve problems, complete a task, or develop a product (Laal & Laal, 2012). In a CLS environment, the groups exploit each other's resources, ideas, skills, and talents. Instructors perform a facilitator's role by offering learners support for determining creative, constructive, and ethical solutions to problems (Klimovienė et al., 2006). The Cooperative Learning process challenges learners cognitively as they engage in constructive listening and share and justify their thoughts without relying on professional opinion or textbook material. A well-structured cooperative learning process can provide understudies with the fundamentals to advance their critical thinking (Cooper, 1995). However, instructors and other learners must provide constant reinforcement to help foster critical thinking skills and dispositions.

Case Studies and Discussions

Case studies and discussion center on active learning; the instructor serves as a facilitator by presenting a topic, scenario, or problem to the class. However, no conclusion is provided (McDade, 1995). The instructor poses questions to elicit dialogue and debate between students, culminating in constructing conclusions. As students share and flow ideas among themselves, they practice inductive learning while building critical thinking capacity.

Socratic Questioning

Socratic questioning is one of the most popular types of questioning. This method involves the instructor posing specific questions to students to promote independent thinking and thoughtful dialogue.

Reciprocal Peer Questioning and Reader's Questioning

The reciprocal peer questioning method involves the instructor presenting students with questions after participating in the lectures. Students also develop questions for small-group sessions and class discussions (King, 1995). Reader's questioning involves students submitting questions from reading assignments for discussions before the class.

Conference Style Learning

CSL involves the instructor facilitating the conference while the learners ask and discuss each other's questions. The learning material allows learners to develop independently while drawing from instructors to create more significant challenges with ideas through strategic assistance or guidance (Underwood & Wald, 1995). For the learning experience to be meaningful, learners must read all necessary material extensively before class. For the learning experience to be significant, learners must read all the required material extensively before class.

Writing

The application of writing is a key element in improving critical thought. Writing helps improve learners' critical thinking skills by encouraging them to recognize issues and propose theories and arguments (Gocsik, 2002). Writing compels learners to clarify their thoughts, connect ideas to formulate logical premises, and present a critical opinion.

Ambiguity

Ambiguity is a teaching approach that provokes learners to better understand the situation or problem before launching a solution or response. Strohm and Baukus (1995) suggested that instructors should produce substantial ambiguity in the classroom by giving learners conflicting and ambiguous information to think through. This approach allows learners to make connections, devise new concepts, and find new meanings.

Many factors interact to determine the effectiveness of developing critical thinking skills. This section addressed methods and formats commonly used to teach critical thinking skills. Knowing these elements helped to shed light on possible methodological approaches that might be used in the various individual studies.

Barriers to Critical Thinking in the U.S. Army

As the U.S. Army's strategic vision and leadership doctrine exalt the value of improving warfighters' critical thinking skills, the assumption is that the PME delivers on this demand (CJCS, 2019; HQDA, 2015; Perkins, 2017). Furthermore, critical thinking is mentioned more than twenty-five times as a fundamental competence or talent to acquire in the U.S. Army. Regardless of the interest in improving warfighters' critical thinking skills, the focus does not appear to align with the practice and culture of the U.S. Army (Reidal, 2002). This section highlights four barriers within the U.S. Army institution that impede the development of critical thinking skills: hierarchical structure and internal cultural norms, ethnocentric and egocentric culture, time and resources devoted to learning and practicing critical thinking, and lack of PME instructors' understanding of critical thinking and quality of instruction.

Hierarchical Structure and Internal Cultural Norms

First, information flows from the bottom to the top in a hierarchically organized organization, with minimal downflow to lower organizational levels, restricting debate and ideas (Senge, 2006). Rank, internal order, and superior conformity that harmonize with hierarchically organized institutions often make it difficult for subordinates to dialogue, raise questions, and offer dissenting opinions (Allen & Gerras, 2009). The principles that govern good thinking are incompatible with a culture that suppresses openness and genuine dissent with superiors. Furthermore, an environment that does not allow transparency and dissent can foster groupthink. While groupthink facilitates cohesiveness and information, it can cause U.S. Army leaders to ignore important alternative options (Riedel, 2002). Groupthink is directive, abandoning others' insights and producing weak decision-making (Chen et al. 1996).

Ethnocentric and Egocentric Culture

The second barrier the U.S. Army faces is its ethnocentric and egocentric culture. Across the U.S., Army forces have adopted an ethnocentric and egocentric attitude that their way is superior or surpasses other military powers through the common belief that they rule the world of warfighting (Monat & Gannon, 2017). The consequences of this mindset prevent the U.S. armed forces from viewing the world from new perspectives, limiting receptivity to better ways of doing things and refusing to change attitudes when appropriate (Allen & Gerras, 2009; Paul & Elder, 2007). White papers produced by the U.S. Army's Human Dimension Capabilities Task Force on cross-cultural competence expressed a strong need to improve on this line of thinking. They appealed to the force to develop a nuanced appreciation for the social context of its operational environments (Brown, 2014).

Devoting Time and Resources to Developing Critical Thinking

Third, leaders across the Department of Defense proclaim the importance of critical thinking skills, but less time and resources are devoted to warfighters' learning and practicing the craft (Murray, 2014). While scholars suggested that it takes three to five hours a week over two years in a college setting to infuse learners with critical thinking skills, the PME curriculum contains 700 hours (about four weeks) of core and advanced-level instruction (Williams, 2013). However, educators only allocate roughly 100 hours (about four days) to critical thinking skills. Consequently, less time is devoted to critical thinking, and students are often overloaded with instructional materials and information to learn in a brief period (Allen & Gerras, 2009). Warfighters not only have to learn critical thinking concepts but also need the time and resources to apply and practice the principles facilitated by an experienced facilitator (Gelder, 2005). According to Gelder, understudies will not automatically develop critical thinking abilities by merely examining a subject or ignoring educational material. In summary, the U.S. Army must place a high priority on critical thinking for it to flourish. This takes both learning the trade and some deliberate practice.

PME Instructors' Understanding and Instruction Quality

The fourth barrier with which the Army grapples in developing soldiers' critical thinking is the uncertainty of PME instructors' robust understanding and instruction quality in implementing critical thinking. To develop critical thinking skills, instructors must be knowledgeable or possess a rich understanding of the role of thinking, intellectual engagement, and requisite teaching strategies (Paul & Elder, 2007; Stone, 2017). Aside from an expectation that instructors' skills should be higher than their students', there is no evidence of defined PME standards teaching critical thinking skills (Allen & Gerras, 2009). It is unclear if PME instructors have the training or expertise to guide and assess students' critical thinking. Allen and Gerras contended that the U.S. Army would benefit from attracting educators with the education, intelligence, and skills required to foster critical thinking. Additionally, the authors challenged the U.S. Army to develop a critical thinking course specifically for the instructors. The foremost authorities on critical thinking, Paul and Elder (2007), also held that teaching critical thinking involves an awareness of the connections between learning, education, and critical thinking. In addition, they state that teaching deep thinking requires a clear conception of critical thinking and the intellectual workings of the mind.

Critical thinking is a fundamental capacity for how warfighters can capitalize on deterrence, fight, and win battles through multi-domain operations (DA, 2018a). Answering the research question depends upon seeing all aspects of the critical thinking capability internal to the U.S. Army. Acknowledging barriers helps to identify areas that potentially impinge upon the U.S. Army's ability to develop soldiers' critical thinking skills.

Relevance of Critical Thinking Assessment

Attending to the relevance of critical thinking assessment is vital for warfighters expected to apply cognitive abilities beyond routine mental activity, such as recall or memorization. Critical thinking assessments evaluate an individual's strength in this area. They are a key element in determining whether warfighters will benefit from critical thinking training or education and their potential to practice the skill in the field. Vogler (2002) asserted that using appropriate assessment techniques facilitates monitoring development and measuring the success of critical thinking applications. A report by the Spellings Commission (2006) noted a lack of quality assurance in higher education that stems from the failure to adequately assess student progress in the outcomes that are deemed most important, such as developing critical thinking skills. In summary, assessments are essential to eliciting evidence of warfighters' growth in critical thinking skills. It is also a tool for providing valuable feedback on instructional outcomes to warfighters and instructors.

Common Critical Thinking Assessments

A range of commercial evaluation instruments aims to measure general critical thinking capacity. The tools' cost-effectiveness and suitability for the Army's specific features, such as the wide range of occupational specialties and career paths and its technically focused, command-centric, and doctrine-based culture, are still being determined (Jackson et al., 2000). This section aimed to demonstrate some of the basic concepts underpinning critical thinking measurement instruments and how they connect to or differ from one another rather than to offer a complete overview of all existing tools. The most widespread critical thinking assessment tools include the Watson-Glaser Critical Thinking Assessment (WGCTA) (Watson & Glaser, 1980), the Cornell Critical Thinking Test (CCTT) Level X and Level Z (Ennis, Millman & Tomko, 1985), the California Critical Thinking Skills Test (CCTST) (Facione, 1990a), the Ennis-Weir Critical Thinking Essay Test (EWCTET) (Ennis & Weir, 1985), the Halpern Critical Thinking Assessment (HCTA) (Halpern, 2010), Critical Thinking Assessment Test (CTAT) (Center for Assessment & Improvement of Learning, 2017), and the Collegiate Learning Assessment (Council for Aid to Education, 2017).

WGCTA

The WGCTA is the oldest critical thinking assessment, developed in the early 1940s by Glaser for his doctoral dissertation (Watson, 1980). The Appraisal assesses an individual’s ability to digest and understand situations and information. Although the WGCTA is one of the most popular critical thinking measures, it covers the fewest critical thinking skills. It comprises 80 multiple-choice reading passages introducing problems, statements, arguments, and interpretations that measure the ability to draw inferences, recognize assumptions, evaluate arguments, and use logical interpretation and deductive reasoning. The principal limitation of WGCTA is that it does not thoroughly test critical thinking. The appraisal does not assess knowledge and attitudes, which characterize critical thinking. Secondly, the WGCTA offers a functional assessment of analytic skills such as deduction, yet the test is limited in measuring induction or generalization. The WGCTA is primarily used in the health care industry, particularly in nursing, to assess understudies’ capacity to think critically in medical care settings. The scope of WGCTA used in a military environment was recorded in two studies, one by the U.S. Air Force (Stone, 2017) and the other by a Rand Study for the U.S. Army (Lytell et al., 2017).

The U.S. Air Force study focused on three factors: The state of critical thinking in the Air Force, recommended metrics to gauge critical thinking skills, and determining whether the existing condition of critical thinking skills aligned with the Air Force Future Operating Concept (Stone, 2017). The researcher guided the study toward active-duty Air Force students attending Air Command and Staff College, School for Advanced Air and Space Studies, and Air War College in the Academic Year 2016. The research yielded several findings, and the most significant are: (a) different modalities of thought can occur among students that do not effectively translate to the WGCTA's measurement; (b) the Department of Defense and Air Force senior leaders consistently stressed the importance of critical thinking. Conversely, they rarely issued advanced directives describing the critical thinking skills to hone and how they should be developed and measured; and (c) the top 20 percent of Air Force officers at the field grade officer level were below-average critical thinkers. The Air Force study's findings showed that the scale of the problem set identified in this study spans beyond the U.S. Army. Furthermore, considering the unique Professional Military Education system environment, assessing the diverse mix of critical thinking skills might prove difficult.

The RAND Arroyo Center conducted a study to develop and test a process to assess critical analytic competencies and proficiency of U.S. Army-enlisted all-source intelligence analysts. The study also involved designing a protocol for ongoing evaluation (Lytell et al., 2017). Although the study offered little information about the use of the WGCTA measurement, it did provide the following findings: (a) there is no guarantee that participants will complete the assessment test, which might result in an insufficient sample size; (b) the absence of repercussions for failing to finish the test is one of several factors that may affect participants' desire to complete evaluative assessment or exercises completing the test; (c) participants could only perform the practical exercises with using guides and reference materials, which may have contributed to not completing the WGTCA test. The WGCTA is the most popular measurement for critical thinking. Still, it shares the challenges of other assessments in costs, correlation with the U.S. Army education, training, and operational environment.

CCTST

The CCTST, developed by Peter Facione in the early 1990s, objectively measures core reasoning skills. The CCTST covers more of the critical thinking domain than the other instruments. The CCTST is a dynamic consortium of tests comprising 52 multiple-choice items using real-life situations and general education for various professional fields and educational levels up to post-secondary schools. It measures critical thinking skills elements such as induction and deduction, observation and credibility, definition and assumption identification, and meaning and fallacies (Insight Assessment, 2016). The CCTST is the first instrument to derive construct validity from the Delphi Report's definition of critical thinking (American Philosophical Association, 1990). Construct validity is a method of determining if the measuring technique used to test a given construct is sufficient and rational and allows the researcher to conclude the sample population (Westen & Rosenthal, 2003). Reliability means the measurement instrument is stable or constant (Laerd Dissertation, 2012). When researchers use the same demographics or conditions, the instrument should yield the same (or nearly identical) outcomes from the administration of one to another. The literature review identified one study that linked CCTST to measuring critical thinking skills within the military. The study published in 2010 aimed to measure the critical thinking skills of U.S. Army Medical Department Center and School (AMEDDC&S) instructors at Fort Sam Houston, Texas. The study intended to establish a benchmark or frame of reference for critical thinking skills development and curriculum design for U.S. Army healthcare professionals, as required by regulatory mandates and professional guidelines (Hobaugh, 2010). The study concluded that there are no critical thinking skill standards expected of instructors aside from the understanding that instructors' skills should be higher than those of their students. The study also discovered that neither the teachers nor the students at the AMEDDC&S had taken additional tests to gauge their critical thinking skills. Because of this, it was impossible to compare the study's findings to further research on military medical teachers' critical thinking abilities.

CCTT

The CCTT Level X and Level Z, developed in 1985, is a widely used general critical thinking assessment test for advanced students in secondary education, higher education, and adults (Ennis et al., 1985). The CCTT is a two-series multiple-choice instrument that assesses five critical thinking aspects: deduction and induction, semantics, observation, the credibility of sources, and definition and assumption. Identification Level X is for grades five through 12, and level Z is for grades ten through 12 and college. Level X includes 71 multiple-choice questions that present fact-based passages to evaluate critical thinking skills associated with induction, deduction, credibility, and identification of assumptions. Level Z consists of 52 multiple-choice questions that present short fact-based passages that measure critical thinking skills related to induction, deduction, and credibility; identification of assumptions; semantics; definition; and prediction in planning. Studies using the CCTT to gauge the critical thinking skills of military personnel, particularly those in the U.S. Army, do not seem to exist.

CAT

The CAT is commonly used in professional fields to assess the analysis, conceptualization, and reasoning components of critical thinking (Ennis, 2003). The test presents real-world situations requiring essay responses to evaluate induction, identification of assumptions, deductions, interpretation, and evaluation of arguments. The CAT is unique in that it is designed to help faculty improve the development of student's critical thinking skills (Haynes et al., 2015). In the literature study, the researcher found no research that linked the Critical Thinking Evaluation Test to the U.S. Army in assessing critical thinking abilities.

CLA

The CLA is a standardized test that measures critical thinking, analytic reasoning, problem-solving, and written communication skills. The assessment, which consists of open-ended questions, is administered to students online and controls for incoming academic ability. Rather than testing for specific content knowledge, CLA tests for general skills (Arum & Roksa 2011). One critique of this tool is that it needs instrumental validity in assessing individual success because it relies more on general ability than domain knowledge (Benjamin & Chun, 2003). Based on this literature review, there is no indication that CLA has been used in the U.S. Army's Professional Military Education system to measure warfighters' critical thinking skills.

EWCTET

The EWCTET is a commonly used open-ended test. It is an essay-based assessment test that quantifies the capacity to dissect, assess, and react to arguments and debates in realistic scenarios (Ennis & Weir, 1985). The test helps evaluate an individual's ability to appraise an argument and formulate an equal written response. The test covers a range of critical thinking competencies such as specificity, identifying reason and assumptions, expressing one's position, considering others' perspectives, avoiding equivocation, and extreme skepticism. EWCTET is appropriate for high school and college students. EWCTET has been criticized for its domain-specific nature, the subjectivity of its scoring convention, and its predisposition to skilled writers (Adams et al., 1996). Based on this literature review, no studies demonstrate using EWCTET within the U.S. Army, particularly the PME, to measure warfighters' critical thinking capability.

HTCA

A contemporary critical thinking assessment with an open-ended format is the HCTA (Halpern, 2010). The HCTA comprises 25 open-ended questions that focus on plausible everyday scenarios, accompanied by 25 concrete questions that prod for an explanation to support each answer. The multi-part nature of the questions enables an assessment of specific CT skills (Ku, 2009). The HCTA questions represent five categories of CT application: Hypothesis testing, verbal reasoning, argument analysis skills, applying relevant probability standards, and problem-solving and selecting solutions. In the current literature review, the degree to which HTCA had been used in the U.S. Army's PMES is unknown.

International Critical Thinking Essay Test

The International Critical Thinking Essay Test assesses critical thinking fundamentals for application in any subject (The Foundation for Critical Thinking, 2019).

The goal of the test is twofold. The first goal is to provide a practical approach to pre- and post-test students to determine the extent to which they have learned to think critically within a discipline or subject. The second goal is to provide a test instrument that stimulates faculty to foster students' critical thinking within their field. The essay test comprises two parts: an analysis of a written scenario and an assessment of a written scenario. In the analysis segment, the student must accurately identify the elements of reasoning within the written response. The student must construct a critical analysis and evaluation of the reasoning that supports the written response in the assessment segment. The test is tailored for secondary and higher-education students and fosters close reading and practical writing skills. The extent to which ICTET was employed in the U.S. Army's PMES is unknown in the current literature review.

MCQ

The MCQ format is used in several tests described above. This format is the least preferred because of its limitations. According to Halpern (2003), using the MCQ format turns the evaluation into a test of verbal and quantitative skills instead of critical thinking. Furthermore, instead of exhibiting their ability to logically analyze, evaluate, and find solutions to problems, MCQ encourages test-takers to guess when they do not know the correct answer. (Ku, 2009). Open-ended questions better measure critical thinking skills by enabling test-takers to exhibit whether they spontaneously use a specific CT skill. No evidence in the literature review shows the use of MCQ in the U.S. Army's PMES to measure warfighters' critical thinking skills.

This section highlighted commercially available critical thinking assessment instruments used in multiple research and studies to test individuals' critical thinking skills. According to the literature review, except for the Watson-Glaser Critical Thinking Appraisal and the California Critical Thinking Skills Test, there is no evidence of experience with these instruments in the U.S. Army PMES. The U.S. Army's lack of engagement with these instruments might be attributed to unfamiliarity or that they may be too costly to implement. The literature review revealed similar measurement aspects of critical thinking among the instruments: reasoning, analysis, argument, and evaluation. Also, the literature review revealed that the instruments varied in the scope of the involvement of critical thinking skills. For instance, some instruments relate to decision-making and problem-solving, while others concentrate on writing or metacognition (Ennis, 2003; Ku, 2009; Watson-Glaser, 1980).

Difficulties with Assessing Critical Thinking Skills

For warfighters to develop, practice, and employ critical thinking skills, instructions, and measurement instruments must be designed to enable them to learn and perform beyond the safety nets of the educational and training environments. Warfighters can only become skilled in critical thinking with proper instruction and assessment of their performance. Assessments are essential to identify strengths and weaknesses, determine the effectiveness of instruction and techniques, and gauge development and progression. Assessing the development of critical thinking skills, however, can be challenging. Researchers have pointed out several challenges in assessing students' critical thinking skills and dispositions. Primarily, these challenges are associated with the reliability and validity of learning instruments and measures, independently and accurately evaluating critical thinking skills and dispositions; students' motivation to learn and practice the skill; confusion about overgeneralization or specificity of the critical thinking skill; and faculty involvement.

Scaling Tests

Scaling tests to demonstrate that students gain critical thinking skills that can be transferred to various settings is challenging when measuring critical thinking. For example, Moss and Koziol (1991) conducted a study to analyze scores of interventions intended to measure critical thinking skills in domain-specific courses. Instructors required students to read a passage, support an inference with argumentation, or evaluate an argument as part of the intervention. According to the authors, no single element enhanced performance through tasks that were supposed to be related. Students' abilities to apply certain critical thinking skills to formulate facts, arguments, assumptions, and conclusions did not generalize through the tasks. In each situation or assignment, anomalous and immaterial characteristics became more apparent in students' performance than the general capacity to think critically. Although researchers have distinguished between critical thinking abilities and dispositions, the continuing dispute over whether critical thinking is general or domain-specific makes assessing the skill challenging (Ennis, 2003; Facione, 1990b; Halpern, 1998; & Lai, 2011). For example, the unresolved debate raises the question of whether understudies can transfer critical thinking skills from one topic to another if they demonstrate them in one. Furthermore, would the inability of understudies to share critical thinking skills imply that they require additional schooling in one or both subject areas (Lai, 2011)? Despite the dispute over the nature of critical thinking, in most assessments, the skill is considered generic (Davies, 2013; Ennis, 1989; Moore, 2011). According to Abrami et al. (2008), designing an evaluation instrument that delineates the distinct effects of critical thinking skills and disposition is a tall order.

Student Motivation

Retaining students' motivation while executing a learning outcome is a constant problem. Several studies have demonstrated that students' motivation in test-taking is frequently insufficient for tests that do not significantly affect performance or other educational outcomes (Barry et al., 2010). Students with high motivation perform better on tests than test-takers with low motivation (Cole & Osterlind, 2008; Wise & DeMars, 2005). Because critical thinking involves reflection, substantial effort, and intellectual commitment, low motivation for test-taking might be problematic for valid and impartial measurement of critical thinking. Liu, Bridgeman, and Adler (2012) affirmed that motivation significantly affects how successfully students perform when learning outcomes are assessed in higher education. The authors argued that test-taking motivation is low because the low-stakes assessments typically have no personal consequences for the tested students. Lai (2011) pointed out that emphasizing that tests have consequences may raise test-taking motivation and improve assessment accuracy.

Instructional Relevance and Legitimacy of Tests

Another challenge of designing a standardized critical thinking assessment is instructional relevance in fostering critical thinking skills. Representatives of the

American Association of Colleges and Universities opined that faculty are increasingly. They are concerned about the legitimacy of tests to assess learning outcomes because, often, test developers produce tests in isolation from curriculum and teaching (AAC&U) (AAC&U 2011). For example, while educational institutions recognize the value of critical thinking abilities, few provide courses mainly geared toward promoting critical thinking (Abrami et al., 2008). Similarly, studies have revealed that an institution's faculty must accept ownership of designing learning outcomes even when assessment results in the need to grow students' critical thinking skills. Understanding and incorporating general education's common goals into learning outcome assessments is critical.

Faculty Involvement

Educators widely understand that faculty involvement is key to assessment impact (Paul, 2005). Nevertheless, gaining faculty involvement and support is a top challenge for learning outcome assessments. Critical thinking is an art that necessitates students' understanding of the concept, its application, and its practice. According to Paul, most college faculty lack a substantive concept of critical thinking. Consequently, they do not (and cannot) use it as a core organizer in the instructional design and cannot teach it. Also, there is a perception among educators that assessments are used mainly to demonstrate openness to external stakeholders as opposed to providing data about understudies' acquisition of critical thinking skills (AAC&U, 2011). As a result, the connection between assessment and teaching needs to be more evident to faculty.

Human Scoring

Multiple assessment formats are the preferred approach for assessing critical thinking (Adams et al., 1996; Ku, 2009; Williamson, Xi, & Breyer, 2012). Despite this notion, using constructed-response items poses questions about scoring. An advantage of human scoring is that scorers can comprehend the content, make judgments on its nature, and evaluate critical thinking skills and the accuracy of the assertion (Shermis & Burstein, 2003). Contrarily, human scoring relies upon the decisions of less-than-ideal human beings. Limitations in specialized training, experience in assessing content, and variations in understanding can impact the reliability of human scoring. Additionally, there is a tendency for humans to inject subjectivity into decisions, limiting the consistency and objectivity of scoring. Automated scoring is a feasible option to mitigate concerns about human scoring. A unique quality of automated scoring is its efficiency and consistency. External factors such as deadlines, bias, preconceptions, or stereotypes do not affect automated scoring. These advantages potentially increase the chances of consistency and objectivity.

Identifying the Appropriate Coursework

An added challenge is identifying coursework that requires students to demonstrate critical thinking skills and effort. When students perform a standardized test and are asked to do something outside the course environment, there might be an issue with the amount of effort they put into it. Students will be less inclined to put forth effort if they do not perceive activities or assignments as having value, relevance to their learning, or a connection to their studies (Heft & Scharff, 2017). They also have an aversion to being constantly tested. The lack of student interest and participation will impact assessment. Misalignment of course activities designed to enhance critical thinking abilities is a similar challenge. Specifically, coursework that fails to match the instructors' assignment expectations of the critical thinking concept. For instance, activities offered for coursework do not lend themselves to critical thinking. To mitigate this challenge, instructors must align coursework and assignments with critical thinking skills.

Selection of Key Variables

A broader literature review conducted on numerous topics and strategies for teaching, learning, and assessing critical thinking skills supports the six levels of Bloom's informed taxonomy model. Based on a broad spectrum of cognitive psychologists, curriculum theorists, instructional researchers, educators, and testing and assessment professionals, Bloom's revised taxonomy is viewed as an unambiguous validated model that influences the development of critical thinking skills (Thompson & O'Loughlin, 2015; Zaidi et al., 2017; Zapalska et al., 2018).

Zapalska, McCarty, Young-McLear, & White (2018) attempted to demonstrate that undergraduate management majors in the Business field could enhance critical thinking skills by accomplishing coursework, assignments, and projects through a sequence of gradually challenging thought practices. The authors adopted Bloom's Revised Taxonomy model to structure projects and reports as tasks for students to accomplish to develop critical thinking skills. The authors purported that if students followed the process of 'memorizing, comprehending, relating, applying, examining, assessing, and creating,' which correlates to Bloom's Revised Taxonomy model, they would develop and master critical thinking skills. The authors inferred that students would gradually develop critical thinking skills in a classroom environment that provides projects and assignments that apply the six levels of Bloom's revised taxonomy model. The authors claimed that Bloom's revised taxonomy is the most commonly accepted and used paradigm for cultivating autonomous and critical thought in academic settings.

According to the authors, utilizing Bloom's taxonomy as a teaching tool can assist in balancing assessment and evaluation questions in class, assignments, and readings to ensure that all orders of thought are employed in students' learning. Christy, Sami, and Arumugam (2020) incorporated Bloom's revised taxonomy model in their work to build a prediction model to analyze students' learning skills. The authors' goal was to create a software tool that would show the application and shortcomings of cognitive skills by using questions corresponding to the six levels of Bloom's revised taxonomy model.

Significant terms (verbs) of the questions became independent variables, whereas Bloom's revised taxonomy levels were designated as the dependent or class variable. The experiment's outcomes revealed that the Support Vector Machine (SVM) classifier can effectively classify the RBT levels of questions with an accuracy of 98%. Also, with accuracy values of 0.83 and 0.79, the K-means clustering findings accurately categorized students of different ability levels, including low, good, and medium.

Khadijeh and Mirzaei Rad (2018) investigated the association between critical thinking and listening comprehension across different genders of Iranian primary English Foreign Language learners. The study aimed to improve listening skills among students of different genders so they can read, analyze, and respond to contemporary worldwide dilemmas in education. The researchers used the California Critical Thinking Skills Test and the listening comprehension test to examine the listening comprehension ability of EFL learners using the six levels of Bloom’s revised taxonomy model. Based on the results of two-way ANOVA, the researchers concluded that adopting the six levels of Bloom's revised taxonomy model in educational systems increases the likelihood of critical thinking.

Using Bloom's revised taxonomy model, Morton Colbert-Getz (2017) explored the influence of the flipped classroom on first-year medical students at the University of Utah School of Medicine. The flipped classroom paradigm is an innovative way to improve student-centered learning (Lage et al., 2000; Moravec et al., 2010; Prober and Khan, 2013; Moffett and Mill, 2014). The study investigated whether the discrepancy in prior research findings was linked to the level of cognition (low or high) required to perform satisfactorily on the course assessment.

The course concluded with students taking a 150-question multiple-choice test to assess all academic topics. The course and assessment were utilized to examine if the flipped classroom (FC) compared to the lecture classroom (LC) improved student performance (Morton & Colbert-Getz, 2017). Based on the performance of the FOM final examination anatomy items, the primary study results classified by Bloom's taxonomy were preferred over the FC students who finished FOM in 2013 and LC students who completed FOM in 2014. Overall, the study results aligned with Bloom's revised taxonomy model, which establishes that a student must possess either a high or low degree of cognition to answer a question correctly.

Faravani and Taleb (2020) conducted a study to ascertain whether teachers' use of higher-order questioning techniques had any discernible effects on Iranian EFL students' capacity for listening comprehension. The researchers also investigated the impact of teachers using higher-order questioning techniques on Iranian EFL students' capacity for critical thought. The participants were college students pursuing various majors or university graduates holding multiple degrees. The first stage of the study consisted of administering two instruments to the participants: a listening comprehension test of TOEFL (TOEFL PBT) and a California critical thinking skill test by Facione (1990).

The researchers administered an English listening comprehension test of TOEFL (TOEFL PBT) twice as a pretest-posttest. The California Critical Thinking Skill test was then offered as a pretest in the experimental and control groups to assess students' critical thinking skills at the start of the classes. The intervention was then administered to the experimental group in the second stage, with students being asked higher-order thinking questions based on ‘Bloom’s revised taxonomy model’ by listening to the book audio Basic Tactics for Listening, 2nd Ed. The final step was to administer post-tests to both groups. To assess the aptitude for listening, students in both groups were given the TOEFL listening comprehension test (TOEFL).

The study findings unveiled that teachers who employed higher-order questions positively impacted learners’ listening comprehension ability. Also, the findings signaled that teachers need to use higher-order questions more than lower-order questions in classes. Akram and Ensie (2020) asserted that one of the most popular hierarchies that might be used in the course is using higher cognitive level questions based on Bloom's Taxonomy.

Qasrawi and Beni Abdelrahman (2020) conducted a quantitative, qualitative content analysis to investigate the extent to which *Intermediate Unlock English Reading, Writing, and Critical Thinking Skills* textbooks enhanced higher-order thinking skills. The researchers analyzed the frequency of the lower and higher thinking learning objectives by creating a checklist with the verbs under Bloom's revised taxonomy model. The authors calculated the number of learning objectives for each activity. The goals were coded and categorized according to Bloom's levels to document the presence of the thinking skills levels (HOTS and LOTS) in the textbooks.

The data analysis revealed that comprehension, a lower-order thinking skill, is where half of the cognitive objectives fall (Qasrawi & Beni Abdelrahman, 2020). The second half of the examined cognitive objectives ranked under the higher-order thinking abilities of analysis and *synthesis*. The analysis concluded that most cognitive objectives belong to LOTS (comprehension) and HOTS (analysis and synthesis). The analysis also revealed that many cognitive goals fell under LOTS (comprehension) and HOTS (analysis and synthesis).

Mulcar and Shwedel (2017) introduced the Critical Reading Topics (CRT) method to determine whether it promotes deep thinking through students' critical engagement, students' active analysis, pointed feedback, and quick assessment. The CRT method consisted of four distinct multistep phases with tasks for students and instructors.

Phase I consisted of an at-home comprehension assignment requiring students to critically read a select set of materials and construct the author's argument before taking a computerized multiple-choice quiz. In Phase II, students performed an in-class comprehension assignment. The students analyzed the lecture and textual material to comprehend concepts and highlighted the most confusing part of the reading. Phase III involved formal grading and instructor feedback. In Phase IV, students received another reading assignment requiring them to read topics, categorize, and raise the text’s main points. In this phase, students could actively interact with the material, learn more about the various levels of Bloom's revised taxonomy, and exercise critical thinking.

The CRT included four levels of Bloom's Taxonomy as expressions of higher-order thinking: analysis, application, evaluation, and creation. According to Mulcare and Shwedel (2017), these four levels of Bloom's revised taxonomy were included in the CRT because they encouraged critical thinking by allowing students to map out the course themes, make connections between readings and other texts, investigate the application of ideas in various contexts, evaluate the viability of an author's arguments, or use concepts to develop novel approaches to the matter.

The studies above are a demonstration of the widespread application of Bloom's revised taxonomy model to various investigations or exploratory studies to teach, learn, or gauge thinking skills (Akram & Ensie, 2020; Christy et al., 2020; Faravani & Taleb, 2020; Khadijeh & Mirzaei Rad, 2018; Mulcare & Shwedel, 2017; and Qasrawi & Beni Abdelrahman, 2020). Bloom’s revised taxonomy model is often used as a guide for categorizing or crafting critical thinking skills questions, checklists, projects, and other activities for learning. More importantly, the literature established acceptance of Bloom’s revised taxonomy as the model to determine objectives, activities, and critical thinking skills assessment.

Conclusion

Developing warfighters' critical thinking skills is a key priority for the U.S. to enhance their cognitive readiness to better execute complex, fluid, unprecedented mission demands. Social science and educational researchers have contributed to defining the concept, development, testing, and measurement of critical thinking skills in many ways. This chapter included a thorough literature review of contemporary theories and practices and methodological advances in assessing and evaluating critical thinking skills. I aspire that this body of literature lays the groundwork for identifying core higher-level thinking features to develop a structured and holistic approach to accessing critical thinking capabilities within the U.S. Army.

Senior military officials hold critical thinking skills as an indispensable core capability that warfighting professionals need to respond to the rigors of great power competition, advanced adversarial threats, and complex challenges worldwide. Despite this, leadership does not know whether soldiers have sufficiently developed their critical thinking skills to apply them in various operating environments. The U.S. Army lacks a single, generally acceptable, and scalable codified tool to assess its force’s critical thinking capabilities. This study can advance the U.S. Army’s warfighting capabilities and literature by demonstrating assessment and measurement methodologies to improve students’ learning and critical thinking practices.

In Chapter 3, I provide a detailed discussion of the methodology for this quantitative meta-analysis study. This includes a detailed discussion of the research design and rationale. I explain the data collection strategy, including sampling and screening. I discuss the data analysis strategy, including statistical tests and the process for interpreting results. A review of validity and ethical procedures concludes the chapter.

Chapter 3: Research Method

I examined common critical thinking skills involving developing reliable tools to gauge U.S. Army warfighters' critical thinking abilities. I used a quantitative five-step meta-analysis design. I examined empirical studies regarding critical thinking skills assessments and measurements and Bloom's revised taxonomy model. I selected studies based on preset criteria to include in the meta-analysis. Then, I extracted and standardized data from selected studies. I calculated the effect sizes from each study. I calculated a combined intervention effect to determine the extent to which the high-level thinking features contribute to developing junior and mid-level learners' critical thinking skills.

The research will serve the DA, PME faculty, TRADOC, and vested educators in terms of enhancing warfighters' cognitive capacities. In addition, the study could promote and provide more detailed answers to improve current critical thinking assessment models and methods. This chapter includes information about the research design, quantitative approach, and meta-analysis design. The chapter consists of the rationale for the design, sampling procedures, data collection procedures and instruments, and data analysis processes and techniques. Possible threats to validity and ethical considerations conclude Chapter 3.

The study involves identifying common critical thinking skills features that might be useful for the U.S. Army to measure warfighters' critical thinking skills and learning achievement regardless of training construct, occupational specialty, subjects, or educational levels. The research question is:

RQ: To what extent can standard features of higher-level thinking (remembering, understanding, applying, analyzing, evaluating, and creating) contribute to developing junior and mid-level learners' critical thinking skills in postsecondary education?

$H_0$1: Standard features of higher-level thinking (remembering, understanding, applying, analyzing, evaluating, and creating) do not contribute to developing critical thinking skills in postsecondary education.

H_a1: Standard features of higher-level thinking (remembering, understanding, applying, analyzing, evaluating, and creating) contribute to the development of critical thinking skills in postsecondary education.

The results of these questions are critical to providing the U.S. Army with a baseline for establishing measurement options for assessing critical thinking skills. Also, anticipated results are expected to improve critical thinking instruction.

Research Design and Rationale

This section includes information about the research design for the study. I adopted the quantitative meta-analysis design to gain a comprehensive coverage of common features that contribute to assessing the development of critical thinking skills. I aimed to provide the U.S. Army with a baseline for establishing critical thinking skills measurement options. Meta-analysis is used for various studies, including examining performance tests, comparing interventions, and assessing program effectiveness. While traditional research methods involve statistical significance testing, meta-analysis addresses the direction and scale of effects across studies (Wilson, 2010).

The general steps of the meta-analysis are: formulating the problem, establishing criteria, searching and selecting studies, assessing risk, extracting effect sizes, coding themes and features, synthesizing effect sizes, analyzing data, and interpreting and reporting outcomes (Basu, 2014). A systematic analysis must precede a meta-analysis to establish that enough data exists from each study. A systematic review is a set of procedures for converting results from several studies to a standard measure (Borenstein et al., 2009). Statistical correlations between variables and outcomes are among systematic review and meta-analysis processes.

Alignment of the problem statement, purpose, and research question supported the application of a quantitative meta-analysis design rather than a qualitative review. The quantitative meta-analysis, which involves using numbers to analyze literature, enables an objective view across a series of studies to explore tools, interventions, diagnostic procedures, and measurements to assess critical thinking skills. The design favors discovering broad patterns across studies, practices, specific interventions, conditions, and treatments (Page et al., 2021; Salters-Pedneault, 2020).

Methodology

Methods of the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement guided the conduct of this meta-analysis. The PRISMA statement is intended to aid in promoting a high-quality meta-analysis. The Cochrane Handbook for Systematic Reviews of Interventions supplemented implementation of the PRISMA statement. The Cochrane Handbook provides rigorous methods for producing high-quality, relevant, accessible systematic reviews and other synthesized evidence.

Incorporating Bloom's revised taxonomy as a criterion for evaluating measurement properties in studies helped strengthen this methodology. Bloom's revised taxonomy is a validated and well-accepted tool that provides a generally acceptable inventory of qualities and characteristics for teaching and assessing critical thinking (Facione, 1990; Lau, 2018). I aimed to develop a baseline for establishing critical thinking skills measurement options based on statistical results using pooled data from single studies to evaluate and calculate critical thinking skills growth.

Population

This research's target population or unit of analysis was primary studies collected from a literature review. A comprehensive literature search was used to identify studies eligible for evaluating critical thinking measurement properties using the (PRISMA statement and Cochrane Handbook to conduct a quality review. The studies included original published and unpublished research articles.

Sampling and Sampling Procedures

This study covered a wide range of studies that were relevant to the research question. Selection of studies was made using the purposeful sampling method. Purposeful sampling is used when particular conditions, persons, or activities are selected deliberately to provide meaningful and rich information that cannot be obtained effectively from other means (Taherdoost, 2016). Purposeful sampling is the most suitable method for conducting research in conditions that meet particular criteria or possess specific characteristics. Using defined inclusion and exclusion criteria specified in the data collection section of this chapter, deliberate sampling consists of original research articles, systematic reviews, and unpublished studies that demonstrated interventions using high-level thinking classifications that were identified via Bloom's revised taxonomy.

Defining Variables

A key phase in the data collection and analysis is defining the variables and considering how they may be applied to explore the research subject. The variables identified for this study are the six high-level thinking features extracted from Bloom's revised taxonomy model: remembering, understanding, applying, analyzing, evaluating, and creating. These higher-level cognitive features are a tiered classification of distinct learning, understanding, and thinking (Lederman, 2017). The higher-level characteristics, which vary from basic knowledge and understanding to higher evaluation and creativity, promote the development of critical thinking and problem-solving skills depending on their complexity. Bloom's revised taxonomy model is a prominent and validated theory in education, psychology, and social science research that provides a systematic classification of the learning and thinking processes to teach, develop, and assess critical thinking skills (Anderson & Krathwohl, 2001; Zapalska et al., 2018). According to Anderson and Krathwohl (2001); Thompson and O'Loughlin (2015); and Zaidi et al. (2017), Bloom's revised taxonomy model exhibits the highest levels of critical thinking, evaluating, and creating. The six levels of knowledge are remembering, understanding, applying, analyzing, evaluating, and creating. Remembering signifies recalling details like dates, events, and concepts. Understanding involves comprehending information, articulating it in one's own words, and providing examples. Applying is using knowledge and skill in new situations. Analyzing involves deconstructing information into its components and explaining their relationships. Evaluating is judging or assessing the value of resources and techniques in specific objectives. Creating refers to forming connections for new situations and consolidating information to create a new whole (Anderson & Krathwohl, 2001).

Bloom's revised taxonomy model is a commonly used framework to design and structure instruction and assessment tools to measure cognitive skills, particularly critical thinking. The six levels of Bloom's updated taxonomy model were the variables examined to understand the hypotheses better.

Data Collection

A meta-analysis requires the specification of inclusion and exclusion criteria. Lipsey and Wilson (2000) stressed that assigning explicit inclusion and exclusion criteria is a characteristic of a good review. Defining these criteria helps to communicate the study's focus and leads to including and dismissing primary research studies. A study had to meet preset criteria based on the primary question under investigation to be eligible for inclusion in this meta-analysis. The study had to be widely accessible, published, or unpublished, or archived. The study had to focus on developing, applying, assessing, and practicing critical thinking. The study must have involved experiments or interventions to improve, measure, or assess critical thinking. At least two independent samples, interventions, or pre-and post-test experimental or quasi-experimental designs had to be used in the study. To extract effect sizes, sufficient statistical data and sample sizes greater than ten participants were required (Hedges and Olkin, (1985).

A meta-analysis literature review was beneficial in identifying empirical studies relevant to the research question. The primary keywords used to identify empirical studies in the literature review included critical thinking assessment, critical thinking measures, teaching critical thinking, and their variations. Also, the literature review involved a search for any insertion of the cognitive skills signified in Bloom's Revised Taxonomy and the Delphi Report. The following electronic databases represent some sources explored: ProQuest Dissertations & Theses Global, Educational Resources Information Center (ERIC); PsycINFO; Academia Social Sciences Index; Military, Education, Philosophy and Psychology journals; and Google Web and Scholar. These databases contain a collection of peer-reviewed studies that expanded access to the sources required to accomplish this meta-analysis.

Power Analysis

No minimum sample size criteria exist to test a hypothesis in a meta-analysis (Turner et al., 2013; Valentine et al., 2010). However, studies conducted in the Cochrane Reviews usually included data from one or more small studies. Given that there is no established sample size benchmark, this study incorporated a power analysis using the Power Calculation for Meta-Analysis tool. I used the standard criterion for detecting a medium statistical power ($\beta = .80$) alpha ($\alpha = .05$) (Cohen, 1988). I also considered Jackson and Turner's (2017) proposed principle, according to which random-effects meta-analysis studies require at least five studies to maintain statistical power. The authors conducted a study to explore methods for evaluating the power of random-effects meta-analyses. The authors' methods are referenced in statistical primers and other published literature. Also, I applied the medium effect size (0.05) and the ideal minimum sample size of 60 (30 per group) for a meta-analysis (Cohen, 1988; Fisher, 1925).

Meta-Analysis Model

The meta-analysis comprises two parts. The first involved calculating a summary statistic and standard of error for each included study to convert the many studies into one standard metric. The second consisted of combining, averaging, and calculating the inverse variance weight for each effect size. The inverse variance helps account for studies that vary in size to better estimate the effect size (Lipsey and Wilson, 2000). The meta-analysis addressed the presence or absence of heterogeneity and variability across studies. It was logical to predict that the studies selected for this meta-analysis would have a different effect because various researchers conducted them, varied in their locations, length, and demography, and used different types of interventions. The random-effects model, which assumes that the effect size might differ from study to study, informed how to address this assumption and made it possible to examine the effect of individual-specific characteristics on a response variable (Higgins et al., 2021). The standard error of the summary represents the confidence interval, which communicates how well the sample data represents the population of the included studies. The standard error of summary also aided in deriving a p-value, which displays the strength of the evidence against the intervention effect.

One of the many practical factors that must be considered when conducting a meta-analysis is missing data (Chilcott et al., 2003; Higgins et al., 2021). Even though the implications of missing data have received most of the attention in the literature, aside from specific statistical tests, there needs to be more guidance on addressing it (Pigott, 2019; Tierney & Cook, 2018). Deeks et al., nevertheless, presented four general recommendations: contact the initial investigators to request any missing data; make predictions on the methods employed to address the missing data; assess the sensitivity of the results to acceptable modifications in the assumptions by performing a sensitivity analysis; explain any probable missing data in the findings of Discussion Section. Other options for dealing with missing data include: limiting analysis to the available data; and substituting values for missing data to treat them as though they were observed. I considered these recommendations and options in addressing missing data, primarily advancing the search to obtain additional information and ensuring to identify and include the impact in the discussion section of this meta-analysis study.

Using the search strategy Participant or Population or Problem, Intervention, Comparison, and Outcomes (PICO), the meta-analysis started with a comprehensive literature review search for relevant critical thinking measurement or assessment studies that evaluate at least one cognitive skill and affective dispositions signified in Bloom's revised taxonomy (Higgins et al., 2021; Richardson et al., 1995). The literature search included a critical appraisal and data extraction of studies that meet the inclusion criteria. The search involved a review of unpublished or gray literature for relevant sources such as reports, dissertations, theses, databases, google searches, studies from other countries, and conference abstracts.

The search also included a check for any relevant retraction statements and errata in information. A comprehensive search is necessary to minimize the risk of publication bias and to uncover as much reliable information as possible (Higgins et al., 2021). Although no formal approaches exist that prescribe when to stop a search, this study adopted the method highlighted by (Chilcott et al.,2003). The search ceased when additional terms introduced to the database search generated no new relevant records, removing keywords or concepts about missing pertinent records or an appropriate level of evidence that had already been gathered.

As recommended by Field and Gillett (2010), I organized and coded studies into categories such as solid or weak, interventions, experiments, and with or without controls. The stated method assisted in deciding whether enough robust studies existed to follow through with the meta-analysis.

A key component of the literature review process is documentation and reporting to support transparency, assessment, and reference for future updates (Rader et al., 2014). Documentation and reporting described the sources searched, when, by whom, using which terms, a summary of the correspondence, and any other information to allow the search to be reproducible. The reporting included each study's basic characteristics, including details of participants, interventions, experiments, control groups, outcomes, and study design (Higgins et al., 2021).

Extracting the Effect Size

Conducting a meta-analysis requires determining the effect size and its standard error. Effect size is a widely accepted measure among education researchers and psychologists to assess the size of differences between two groups (*Statistics Solutions,* 2013; Waheed, 2020). Effect size is essential because it tells the magnitude and direction of the relationship or the strength of the relationship between two continuous variables (Field & Gillett, 2010). A meta-analysis is used to aggregate the effect sizes of several related studies to determine the average effect size of a particular finding. Large effect sizes indicate that the difference is important, while small effect sizes suggest that the difference is unimportant (Bernhardt, 2004; Coe, 2002; Madsen et al., 2016). The individual studies yielded one or more effect sizes. Due to the relationships between effect sizes, extrapolating multiple effect sizes from the individual study raises the risk of biases. It also conflicts with the study's objective of combining and averaging the various studies into a single standard measure.

Several approaches measure the effect size: the Pearson product-moment correlation coefficient, *r*; the effect-size index, *d*; and odds ratios, risk rates, and risk differences (Law, Schmidt, & Hunter, 1994; Statistics Solutions, 2013). Notwithstanding, there are various proposed variations to the metrics, such as Glass's *a*, Cohen's *d*, and Hedges's *g*. It is important to be familiar with the type of data (e.g., dichotomous, continuous) that results from measuring an outcome in an individual study and choosing suitable effect measures for comparing intervention groups' differences. Since Cohen's *d*, Pearson correlation coefficient *r*, and the odds ratio (OR) are the most often used in social sciences, these approaches would apply to the present study as required.

This research used Cohen's *d* or standardized mean effect to compare the means from studies with group designs and continuous outcome measures. Studies using treatment or control groups are characterized as group designs. According to Cohen's (1992) broad criteria, there are three forms of effects: large, medium, and small effects: *r*

14:10 (small effect, accounting for 1% of the total variance); *r* 14:30 (medium effect), which accounts for 9% of the total variance; and *r* 14:50 (large effect), which accounts for 50% of the total variance. It is important to note that the above guidelines are notional for use as an initial basis for determining the value of an effect (Baguley, 2009; Field & Gillett, 2010). Consequently, they might not always be comparable when converted to different metrics. However, presently, no practical alternative exists for evaluating an effect size within the context of a research domain.

Calculating Effect Size and Standard of Error

To normalize Cohen's *d* effect size for this study, the means of the two groups' means had to be subtracted, and the total of the squared errors had to be divided: The formula used was:

$$d = \frac{\text{m1 (group or treatment 1)} - \text{m2 (group or treatment 2)}}{\text{[pooled] sd}}$$

Nonreported Statistics

While conducting the meta-analysis, there were instances where studies did not report statistics such as mean, standard deviation, and t-values needed to calculate effect sizes (Waheed, 2020). In cases of non-reported statistics, I relied on any of Waheed's proposed coding solutions. I categorized the effect size as zero for insignificant results and unreported data. I coded the effect size with the corresponding p-value if the results were significant and the p-value was provided. I estimated the effect size using the minimal p-value threshold of 0.05 when the results were significant, but the p-value was not provided.

Transforming Effect Size

Once the appropriate effect size has been extracted from the studies, it must be transformed to balance any bias in overestimating effect sizes. Cohen's *d* is the statistic to transform the effect size. However, Cohen's *d* tends to overstate effect sizes in small samples. Hedge's *g* is an added statistic to correct upward bias in sample sizes below 20.

Hedge's formula: $g = (x1 - x2) / \sqrt{((n1-1) *s12 + (n2-1) *s22) / (n1+n2-2)}$

The values extracted using Pearson's *r* tend to be skewed. Since values cannot be part of the meta-analysis, they must be transformed. In instances where values existed, the Fisher's *z* statistic would convert Pearson's *r* values into a normal sampling distribution.

Fisher's *z* transformation formula: $z' = .5[\ln(1+r) - \ln(1-r)]$

Standard Error of Effect Size Computation

Computing the standard error is another component of calculating the effect size. The standard error of effect size indicates the estimate's accuracy relative to the mean. The following formulas are used to compute the extract effect size from the studies:

Standard error formula associated with Hedge's *g*: $SEg = \frac{\sqrt{ng1+ ng2}}{\sqrt{ng1\ ng2}} + \frac{\sqrt{g2}}{\sqrt{(2ng1 + ng2)}}$

Standard error formula associated with Fischer's *z*: $SE\ r = \frac{1}{\sqrt{n-3}}$

Inverse Variance Weighting

I also factored in that the effect sizes are not created equal— larger studies with large samples have smaller standards of error and should carry more weight than small studies with fewer samples (Hedges, 1983). Therefore, preference to larger samples received a higher preference by incorporating the inverse variance weighting. According to Hedges, the sample size of individual studies is a preferable weight for the inverse variance method. The weighting was applied to Fisher's *z-effect* sizes.

Weight for Fisher's *z* $\quad Wg = \frac{1}{SE\ /2zr}$

Coding Scheme

When the effect size has been appropriately extracted from studies and calculated by applying the various statistical measures, the effect size values are then identified as the dependent variables of the meta-analysis (Waheed, 2020). Moreover, the study characteristics become independent variables. A critical objective of a quantitative meta-analysis study is to characterize and explain factors that generate variation among studies, such as methods, measures, interventions, and context. Also, it is necessary to distinguish between the phenomenon under investigation and research methods, such as the design and procedures.

This research contained a coding scheme that recorded all necessary information about the research procedures; each included the study and factors that generated variation among the studies. The coding scheme contained summaries of the studies, type of research, intervention, critical thinking assessment variables, sample size, effect sizes, type of measurement, education level, and study discipline or sector of study.

Coder Reliability

Drawing a subsample of the coded study and coding again after a specified amount of time for reproducibility helped to assure the reliability of the coding scheme (Lipsey & Wilson, 2001; Waheed, 2020). According to Lipsey and Wilson, a subset of 20 is sufficient to compare, with 50-plus being the most desirable.

Data Analysis

This meta-analysis's analysis component comprised four elements: type of model, report summary results, publication bias assessment, and moderator analysis.

The meta-analysis follows two models: fixed-effect and random-effect (Borenstein et al., 2011). The fixed-effect model assumes that one true effect size exists for all included studies in the meta-analysis (Cohn & Becker, 2003; Lipsey & Wilson, 2000; Waheed, 2020). In other words, intervention works the same regardless of the population. Another assumption under this model is that any difference in the effect size distribution results from sampling error. When sample effect sizes are drawn from the same population with set average effect sizes, they are predicted to be homogenous. Homogeneity is a common assumption across various statistical analyses, such as the t-test and ANOVA. The fixed-effect model does not require a moderator analysis.

The second model is the random-effects model, which assumes that the effect size might differ from study to study. The term "random" represents that the included studies are a random sample of all studies that meet the inclusion criteria (Waheed, 2020). This model requires a moderator analysis. Each model holds statistical variations. In fixed-effect models, a within-study error results from sampling studies from a population of studies error (Cohn & Becker, 2003; Lipsey & Wilson, 2000). The same error can be present in random-effects models, along with the potential between-study error because of the random sample studies of the general included studies error.

As advised by Borenstein et al. (2011) and Lipsey and Wilson (2000), the random-effects model was implemented in the current study for two reasons. First, in the social sciences field, research is naturally undertaken by a diverse group of researchers employing a wide range of approaches, resulting in a wide variety of effect sizes (Field, 2003, 2005; Hunter & Schmidt, 2000). As assumed, the research differed in instructional techniques, intervention, learning outcomes, and assessments. Secondly, a random-effects model's findings are unconditional and could be generalized to a population of studies greater than the sample. The fixed-effect model is best for drawing conclusions limited to the studies included in the analysis. A random-effects model is ideal because it is unconditional and enables social science researchers to extend their findings to further research.

Study Results

Using the software IBM SPSS Statistics Version 28, the summary of results of the random-effect model included the following: the computed mean effect size or aggregate effect size; associated standard errors; *p*-values; confidence intervals; heterogeneity; cases analyzed; and a forest plot (Higgins et al., 2021; Waheed, 2020).

Forest Plot

A forest plot, a graphical representation of the summary of the study's analysis and results, is the principal outcome of any meta-analysis (Lewis & Clark, 2001). The forest plot displays estimates of effect size and confidence intervals for each study, an overall estimate of effect size, and the accompanying overall confidence interval for all included studies. In addition to illustrating the effect sizes and related confidence intervals of individual studies, the forest plot demonstrates the extent to which the results from individual studies vary.

Publication Biases

Many biases can occur in analyzing meta-analytic data, which must be addressed in the research. The primary bias is publication bias. According to Rosenthal (1979) and other meta-analysis experts, such as Higgins et al. (2021) and Vevea and Woods (2005), publication bias is one of the most severe threats to validity in a meta-analysis study. This means that the meta-analysis has included only a selected category of studies. The current study used the funnel plot to uncover any publication bias. The funnel plot is a straightforward and efficient graphical approach for exploring potential publication bias (Page et al., 2021). A funnel plot displays effect sizes plotted against the sample size, standard error, conditional variance, or some other measure of the estimate's precision. A cloud of symmetric data points around the population effect size with a funnel shape demonstrates an unbiased sample. A funnel shape represents a significant variability in effect sizes between studies with small sample sizes/less precision.

Publication bias is a reasonable possibility in this study because research with small samples indicating small effects is less likely to be published than studies with the same sample size but with more significant effects (Macaskill, Walter, & Irwig, 2001). Since funnel plots and the associated measures do not offer a way to rectify possible bias, a sensitivity analysis would be employed to supplement the funnel plot. Although there does not appear to be a single approach to conducting a sensitivity analysis, the fundamental concept is to repeat the primary analysis while altering the datasets or statistical methods to determine whether the changes affect the combined outcome estimate. In cases of detecting publication bias, I excluded from the analysis studies of lesser quality and repeated the process until the bias was resolved. Minor changes in the overall outcome estimate indicated a sound analysis.

Assessing Heterogeneity

In a meta-analysis, heterogeneity refers to variations in study outcomes across studies. Numerous factors influence heterogeneity, including: (a) the amount of studies used in the meta-analysis; (b) studies' effect sizes variation (between-studies variance); and (c) variation in the observed effect size for the individual study (within-study variance) (Higgins, 2008). The *Q*-statistic is one of the primary statistics used to assess heterogeneity in and between studies. The *Q*-statistic can test the null hypothesis that a single effect size applies to all studies. It is the weighted sum of the squared values of each study’s effect size deviation from the mean effect size of all studies in the meta-analysis.

When testing for heterogeneity using the *Q* statistic, the number of studies in the analysis (k) determines the degrees of freedom ($df = k–1$). If $Q – df$ is less than zero, then there is no heterogeneity beyond what is expected if all studies had the same true effect size. If $Q - df > 0$, then there is significant variance in the effect sizes among studies, which may be attributable to sources of clinical or methodological heterogeneity (Higgins, 2008). The values of *Q* and the *P* value of the *Q*-statistic are dependent on the number of studies in the analysis. The *Q* statistic has a chi-squared distribution, and its value is commonly expressed as chi-squared or χ^2 in summary statistics.

Moderator Analysis

It is expected to conduct a moderator analysis to account for diverse effect sizes and determine if the discrepancies in effect sizes among studies relate to variations in methods, studies, or interventions (Deeks et al., 2022). Meta-regression and analog ANOVA are two main models to explain the heterogeneous effect sizes. Adopting these models requires sufficient study details to allow for a reliable analysis to determine the cause of the variations. Based on the above, a meta-regression and analog ANOVA were not possible in this meta-analysis. Therefore, a moderation analysis was adopted as the most acceptable method to address the conforming variables.

Statistical Software

Most statistical processes and data analysis identified in this methodology required using one or more software programs. I used the IBM SPSS Statistics Version 28 software to accomplish the statistics and analysis requirements.

Summary

This chapter included a descriptive and in-depth discussion of the methods involved in the study. I used a quantitative approach and meta-analysis strategy to examine the extent to which higher-level thinking features from Bloom's revised taxonomy model contribute to developing critical thinking skills. The chapter includes several important and comprehensive steps for executing a meta-analysis, including conducting a thorough literature search of studies that are relevant to the research problem, data collection for aggregating studies and evaluating inclusive criteria, statistical analysis for determining a common effect size, and weighting each study based on sample size to establish a true effect size. In Chapter 4, I present findings and results and any deviations from data collection and analysis procedures described in Chapter 3. Results and findings are presented as descriptive statistics, including the sample and statistical findings. Any new statistical tests that were discovered throughout the research are included in the chapter.

Chapter 4: Results

The CJCS declared critical thinking a capability imperative for Force XXI. The Chairman charged Military Services to provide education and complementary training to produce professionally competent critical thinkers. The U.S. Army depends on its PME system to teach and develop critical thinking skills. The U.S. Army lacks a single, generally accepted, and codified tool for assessing the qualities of warfighters' critical thinking skills. I aimed to address the following research question through a quantitative method using a meta-analysis design:

RQ: To what extent can standard features of higher-level thinking (remembering, understanding, applying, analyzing, evaluating, and creating) contribute to developing junior and mid-level learners' critical thinking skills in postsecondary education?

$H_0$1: Standard features of higher-level thinking (remembering, understanding, applying, analyzing, evaluating, and creating) do not contribute to developing critical thinking skills in postsecondary education.

H_a1: Standard features of higher-level thinking (remembering, understanding, applying, analyzing, evaluating, and creating) contribute to the development of critical thinking skills in postsecondary education.

This chapter includes information about how the meta-analysis process was used and findings based on the study's research question and hypotheses. I describe how data from combined studies were used to create a composite weighted sample, analyze variance within and between studies, estimate the degree of heterogeneity, and then determine the overall population effect size using a random effects statistical model. The statistical model guided the analytical process of determining the extent and magnitude to which common higher-level thinking features contribute to the growth of critical thinking, along with analyzing conforming variables that may have potentially influenced meta-analysis findings.

Data Collection

In a meta-analysis, studies are the principal unit of interest. For a high-quality systematic review and subsequent meta-analysis, it was essential to conduct a thorough search to find a sufficient quantity of eligible studies from potential sources. I researched and evaluated existing studies to determine whether to retrieve them. Choosing which studies to include was based on comprehensive article reviews. Inclusion criteria used to determine eligibility were that they had to be accessible, available publicly, or archived.

Studies had to address the issue of critical thinking development, improvement, application, or practical use. All studies had to employ some defined intervention. It was crucial to select studies that compared results of various types or intensities of treatment, such as studies that included control and experimental groups and pre and post-test results. Studies had to include enough quantitative information, such as measurements of important dependent variables, to allow for analysis of effect sizes. Participants' education levels had to be at the college/university level. Studies were excluded if they did not use pre or post-test experimental or quasi-experimental designs and at least two independent samples. I also rejected sources with insufficient statistical information to extract effect size and studies published before 2008 or not written in English. I also dismissed sources with sample sizes of less than 10 participants.

Selection of Studies

A literature search was conducted to identify and extract empirical studies relevant to the research question. I analyzed the eligibility of articles by screening their titles and abstracts. Subsequently, I reviewed full-text articles for final inclusion. The search was limited to empirical studies written in English that used interventions or experiments to measure or assess the development of critical thinking skills. Peer-reviewed journal articles, conference papers, degree theses, and doctoral dissertations available in online sources were included in the literature search to ensure coverage of empirical studies on the topic in question. All sources were published between 2008 and 2023. I chose this arbitrary period to limit the literature search to the most relevant and recent empirical studies. Bloom's revised taxonomy model was the basis for identifying the most common variables used in studies to measure critical thinking skills. Using keywords and descriptors, I assessed several databases (see Table 1) to locate studies exploring critical thinking features.

Table 1

Databases and Keywords

Database searched	Keywords/Terms searched
EBSCO Professional Development Collection EBSCO Psychology	testing the development of critical thinking skills, assessing critical thinking skills development, Interventions to develop critical thinking skills, critical thinking skills development assessment, critical thinking skills development interventions, critical thinking skills development pre and post-test, testing CT skills development, critical thinking, and measuring outcomes, critical thinking assessment approaches, critical thinking in the military
EBSCO Academic Search Premier	critical thinking pre and post-test, critical thinking skills development interventions, critical thinking in the military, critical thinking measurement instruments, variations in critical thinking assessments
ABI/Inform Global on ProQuest Central	critical thinking skills development interventions, critical thinking in the military, critical thinking measurement instruments, innovative approaches to assessing critical thinking skills
Science Direct	critical thinking assessment test, measuring the development of critical thinking skills, validated critical thinking skills assessment, critical thinking assessment techniques, critical thinking assessment approaches
Sage	Test that measures the development of critical thinking skills, assessing critical thinking skills development, Interventions to develop critical thinking skills, critical thinking skills development assessment, critical thinking skills development interventions, critical thinking skills development pre and post-test, testing CT skills development, critical thinking, and measuring outcomes, critical thinking assessment approaches, critical thinking skill indicators
Behavioral Sciences Collection	developing critical thinking skills

Database searched	Keywords/Terms searched
Google Scholar	Explore the relationship between instructional interventions and CT skills in the setting of postsecondary education, instruments to measure teaching methods that enhance critical thinking skills, critical thinking skills development interventions, critical thinking skill indicators, interventions of developing critical thinking, measuring the development of critical thinking skills, validated critical thinking skills assessment, critical thinking assessment techniques, and critical thinking assessment approaches, professional military education and developing critical thinking
Thoreau Multi-database	Test that measures the development of critical thinking skills, assessing critical thinking skills development, Interventions to develop critical thinking skills, critical thinking skills development assessment, critical thinking skills development interventions, critical thinking skills development pre and post-test, testing CT skills development, critical thinking, and measuring outcomes, critical thinking assessment approaches, critical thinking skill indicators
Social Science Research Network	measuring critical thinking development, critical thinking development
Proquest Central Identifying CT skills	explore the relationship between instructional interventions and CT skills in the setting of postsecondary education, instruments to measure teaching methods that enhance critical thinking skills, critical thinking skills development interventions, Interventions of developing critical thinking, measuring the development of critical thinking skills, validated critical thinking skills assessment, critical thinking assessment techniques, and critical thinking assessment approaches
Cochrane Database of Systematic Reviews	critical thinking skills, developing critical thinking skills
Social Sciences Full Text	critical thinking assessment test, measuring the development of critical thinking skills, validated critical thinking skills assessment, critical thinking assessment techniques, and critical thinking assessment approaches, innovative approaches to assessing critical thinking skills
Education Research Complete	critical thinking skills variables, military and critical thinking skills, critical thinking skills, critical thinking instructional and measurement approaches, validating critical thinking skills

A total of 131,066 studies were discovered in the literature search. The following criteria were used to evaluate whether to include the studies in the meta-analysis:

measured or assessed the development of student's critical thinking skills, conducted in the context of higher-level education, provided sufficient quantitative data to calculate the standardized mean gain effect size for within-group designs and standardized mean difference effect size for between-group designs, and the study used the tests, experiment, questionnaires, or standardized instruments. I closely examined the studies' contents to verify that there was no duplication among the journals, conference publications, and dissertations. I eliminated seven studies because they were duplicates of existing studies. The review resulted in selecting thirteen studies to include in the meta-analysis. Effect sizes were calculated independently for the studies that compared more than one pair of intervention-control groups. The calculation yielded a total of 40 effect sizes. The Characteristics list in Table 2 includes studies, study descriptors, central elements of the intervention, and sample characteristics.

Table 2

Study Characteristics (N = 13)

Study ID	Author/ Year	Research design	Participants	Intervention	Sector of study	Education level	Variables
#01	Esmaeilzad, S. (2022)	Experiment	60	Pre/post-test	Other (librarian)	Undergrad/ Grad-Postgrad	analysis, inference, evaluation, inductive reasoning, and deductive reasoning
#02	Yousef, W. (2021)	Experiment	80	Pre/post-test	General Studies	Undergrad	inference, recognition of assumptions, interpretation, strategy formulation, hypotheses construction, and explanation skill
#03	Lovelace, K. (2016)	Correlational	98	Pre/posttest no control group	Business	Undergrad	identify & summarize problem(s); identify assumptions; analyze reasonable alternatives & consequences;

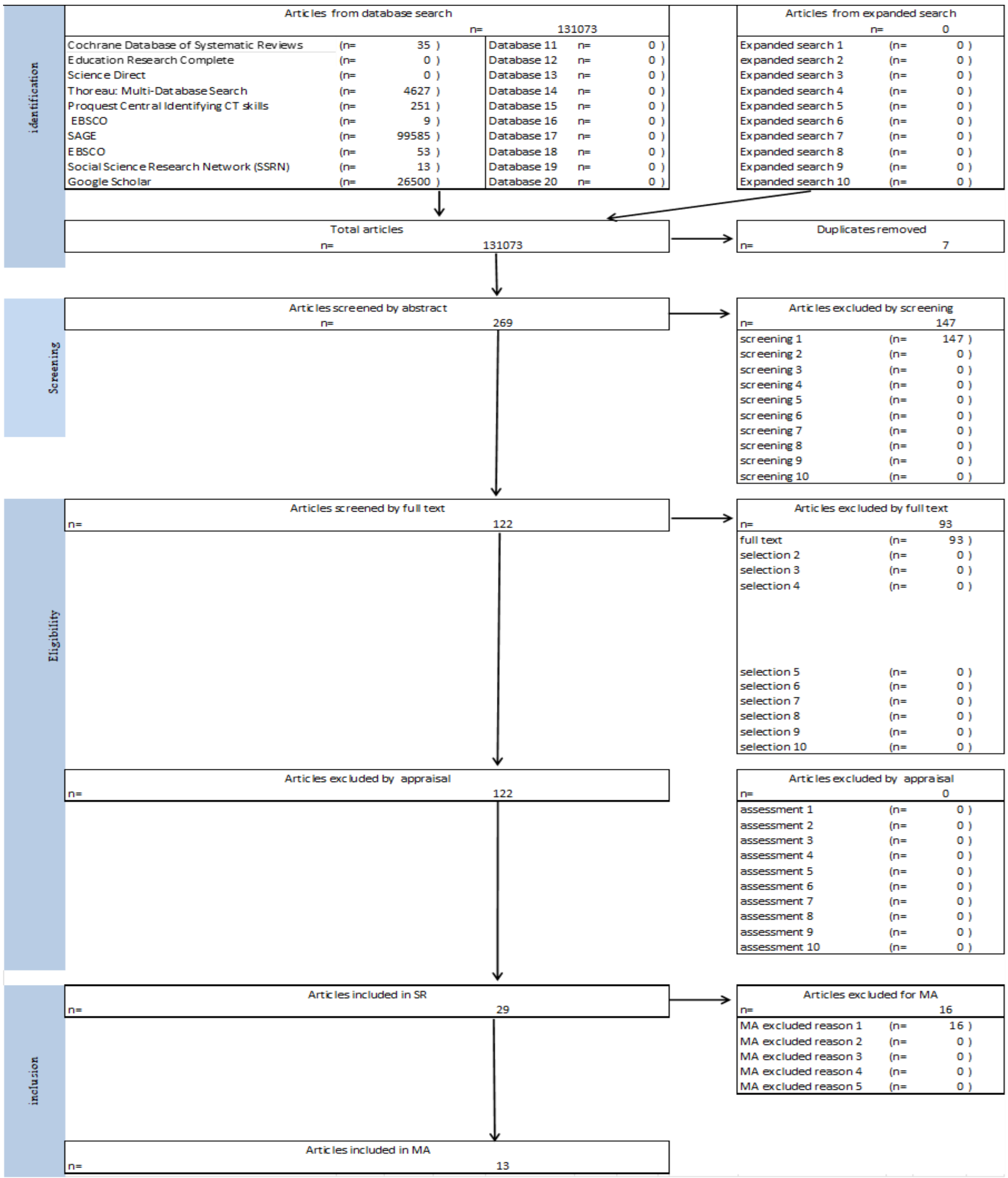
Study ID	Author/ Year	Research design	Participants	Intervention	Sector of study	Education level	Variables
							analyze & present supporting data; provide recommendation & course(s) of action
#04	James, D. (2016)	Longitudinal	44	Separate Groups yrs 1-4	Pharmacy	Undergrad	identifying and considering context, problem-solving, and innovative thinking
#05	Shavelson, R. (2018)	Evidence center design	30	Pre/posttest (separate groups)	General Studies	Undergrad/ Grad	trustworthiness of the information; relevance of the information; proneness to judgmental/decision /bias; reach a judgment, reach a decision; recommend a course of action, suggest a problem solution
#06	Hayes, K. (2008)	Longitudinal	50	Pre/post-test	Food Science	Undergrad	analyze, evaluate, & extend an argument
#07	Vec, T. (2019)	Longitudinal	100	Separate Groups PV/SP (ID Problems)	Socio-Economic/ Education	Undergrad	recognize assumptions, evaluate arguments, distinguish facts, opinions, and conclusions, and make conclusions
#08	Fischer, S. (2009)	Experiment	19	Separate Groups /CT training		Military	identify a frame for a structured message, analyze an the unstructured message, identify weak spots in a message, resolve weak spots, critical self-assessment, extract gist from a message, and find the anchor points in a message
#09	Jimenez J. (2021)	Cross sectional	215	Separate Groups (1st yr/4th yr)	Nursing	Undergrad/ Grad	reading, writing, listening, and speaking
#10	Wallace, E. (2015)	Experiment	76	Pre/Post-test	General Studies	Undergrad	access, communicate, create, define, evaluate, and integrate
#11	Petcharuk, R. (2021)	Research & development	30	Pre/Post-test	Engineering	Undergrad	problem identification, understanding and goal setting, planning, data gathering, experiment,

Study ID	Author/ Year	Research design	Participants	Intervention	Sector of study	Education level	Variables
							conclusion, and evaluation
#12	Jacob, S. (2009)	not provided	40	Pre/Posttest - Clarification	Math/ Engineering	Undergrad	clarification formulates the problem, assessment, raises questions and problems within the problem, inference, reasons, strategies, thinks and suggests open-mindedly
#13	Berry, T. (2013)	Experiment	38	Pre/Post-test	Nursing	Undergrad	explanation, interpretation, inference, evaluation, analysis, and self-regulation

The PRISMA Flowchart in Figure 1 illustrates the process implemented to identify the final articles in the systematic review. A total of 131,073 articles were identified from multiple databases, including a separate review of references contained in the retrieved articles. After eliminating the duplicates, I conducted a general screening of the 131,066 remaining articles. Subsequently, I performed a title and abstract screening of 269 articles retained from the general screening, followed by a more in-depth full-text analysis. Through the analysis, thirteen articles were identified and selected, which met the sampling conditions for this study.

Figure 1

PRISMA Flowchart



Data Extraction

Following the selection of the publications, I coded and analyzed certain aspects of the studies and the outcomes presented. Details on the nature of the research, participants, age range, educational levels, interventions, statistical techniques, discipline or field of study, length of the study, and indicators or traits of critical thinking were among the data that were retrieved.

The included studies were conducted in colleges and universities in various locations: The United States 9 (62%), Spain (.13%), the Middle East 2 (.26%), and Slovenian 1 (.13%). The studies disciplines were composed of general studies (.39%), followed by nursing, education, and mathematics/engineering (.13%) respectively, and others (.39%). The sample size was sixty-eight participants on average, considering the 13 studies included in the review. Jimenez (2021) had the largest sample (*n* =215), whereas Fischer had the smallest sample size (*n* = 19). The participants' ages ranged from 18 to 56.

Coding Procedures

Coding in a meta-analysis serves two purposes. Coding is a technique to highlight the primary study settings, participants, research methodology, and outcome to understand the constraints of the external validity review more clearly (Wood & Eagly, 2009). Also, coding is for examining how the effect size varies across the studies' methodologies, settings, participants, and other characteristics.

Information Retrieval and Coding

After the full texts had been determined to match the criteria for inclusion, I extracted and coded various aspects of each study. I categorized the codes into six areas: study summary, type of research, intervention, critical thinking assessment variables, sample size, type of measurement, education level, and study discipline or sector of study.

The study summary consisted of the author's name, year of the study, type of report, and any information that identifies the study. The purpose of providing the report type was to distinguish whether a study was published. This approach took all feasible measures to eliminate or significantly diminish publication bias, an essential objective in performing a meta-analysis (Borenstein et al., 2021).

Next, I coded for the study design and randomization. Additionally, I coded the duration of the study.

I coded relevant details of the study's settings, age, education levels, discipline, or sector of study. This data made it possible to examine the applicability of studies to their participants and evaluate how well the studies applied to the population (Li et al., 2022).

I coded the indicators or features of critical thinking used in the sampled studies' measurement or assessment approach. Also, if provided, I included the author's description of the indicator or feature.

Next, I coded sufficient information about the intervention, including type, method, mode of administration, name of scale or measurement, reliability of measure (if provided), and any other relevant considerations or components. These details are critical to enabling the replication of the study.

In coding the outcome characteristics, I was particularly interested in how the intervention affected the participants' critical thinking skills.

I independently coded the studies using segments of the Pieces Workbook designed by (Foster, 2018). The workbook was used to guide the data collection and organize and code data from the studies. To verify the reliability and consistency of the coded data, I trained an independent reviewer to random sample the extracted coding sheets and compared them against the respective individual studies. The independent reviewer conducted the review, identified minor consistencies related to transposed numbers, and inadvertently misinterpreted the standard deviation in one data set. I resolved these inconsistencies and resubmitted the updated data sheet to the reviewer. The reviewer did not identify any further discrepancies, which yielded a high confidence level in the accuracy and quality of coding.

Results

The results of this research are explained in multiple sections. The first section presents an overview of the selection and final extraction of the studies included in the meta-analysis. The second section covers the assumptions; the third section presents statistical evidence on the effects of the common features of higher-level thinking on developing junior and mid-level learners' critical thinking skills in post-secondary education. The fourth section explores confounding variables, followed by a summary.

Overall Analysis

An essential step in conducting a meta-analysis is calculating the effect sizes. I used descriptive statistics (means, standard deviation, and sample size) to calculate the effect sizes of the selected studies. I applied Cohen's *d* statistics to measure the mean difference and characterize the effect size. Since it seemed improbable that these studies all share a common treatment effect, I used the random effects model to calculate the effect sizes. The random-effects meta-analysis model considers that sampling variability and actual differences in the treatment effect in each study might cause the observed estimates of treatment effect to vary across studies (Borenstein et al., 2009). In contrast, the fixed effect model assumes that all studies share a common true effect size and that any observed variation is due to sampling error only. Compared to the fixed effect model, the random effects model is more conservative; it distributes weights to studies based on both within- and between-study variation and produces larger confidence intervals. To assess whether there was a significant variance between studies, I examined the Q and I^2 statistics. The Q statistic represents the observed dispersion in effect sizes. It is compared with the value of k-1 degrees of freedom, where k is the number of studies. The value k-1 is the expected Q value, assuming all studies share a common true effect size. Since the Q statistic is sensitive to sample size, the I^2 statistic assesses the proportion of variance between studies attributable to true variance instead of sampling error (Borenstein et al., 2009). The IBM SPSS Software, Version 28, was used to conduct this analysis.

Statistical Assumptions

Statistical tests produce some common assumptions that must be met before the test can be used. The first assumption evaluated was ensuring that effect sizes are independent. This assumption is invalidated if multiple outcomes measure is used for the same participant sample if two treatment groups use the same control group, or if the meta-analysis includes duplicate data or samples (Becker, 2000; Hedges et al., 2010; Liu et al., J., 2020; & Wood, 2008). I verified this assumption by examining the primary studies' methodology to determine whether samples were collected using random sampling. The next assumption I examined was to confirm that the observed estimates of the intervention effect vary across studies because of actual differences in the interventions of each study and sampling variability. Applying the random-effects model establishes that factors exist that contribute to variations in the magnitude of the effect. In the present study, an evaluation of assumptions showed no appearance of a violation.

A search for relevant studies yielded 131,073 distinct articles. After reviewing the titles and abstracts, I selected 122 articles for full-text retrieval and evaluation using the inclusion criteria. The evaluation resulted in a selection of thirteen studies, spanning the years 2006 to 2022, with a total of forty data sets. The total sample size of participants across all studies was 880 (*N*), with an average sample size of 68 per study.

To find the answer to the research question, 'To what extent do the common higher-level features contribute to developing critical thinking skills?', the studies included in this meta-analysis were integrated with Cohens'*d* and the p-values to generate a common effect size.

Effect Sizes

Results of the overall effect size analysis and test for heterogeneity are presented in Table 3, Effect Size Estimates. When all 40 effect sizes are combined, the overall estimate of the effect size of 0.501 was statistically significant (*SE* = .125, *z* = 4.003, $p <$ .001). The width of the confidence interval is dependent upon the correctness of individual studies and the number of cumulative studies (Zientek et al., 2012). The 95% confidence interval of the mean effect size is (0.256, 0.747), indicating that 95% of the time, the true mean effect size of an instructional and training intervention will generally range between 0.256 and 0.747. The results verify that the common features tested by the studies interventions obtained a narrow interval, giving rise to confidence in the level of variance of the mean effect sizes. This means that the intervention outcomes from the combined studies demonstrated that the common features of higher-level thinking: remembering, understanding, applying, analyzing, evaluating, and creating have a moderate effect on the development of junior and mid-level learners' critical thinking skills.

Table 3

Effect Size Estimates

Effect size	Std. error	*Z*	Sig. (2-tailed)	95% confidence interval		95% prediction interval	
				Lower	Upper	Lower	Upper
0.501	0.1252	4.003	0.000	0.256	0.747	-1.018	2.020

a. Based on t-distribution.

The test for heterogeneity yielded a *Q* statistic of 289.931 compared to an expected value of 39. The test of the null hypothesis is statistically significant ($p < .001$). The I^2 statistic estimation is 91.0, meaning that 91.0% of the observed total variance between studies is due to fundamental differences in the effect size. The variance of the mean effect size is 0.547 (T^2 statistic), and the standard deviation is 0.74 (*T* statistic). The statistically significant variability of the effect sizes implies that the finding is incongruent, requiring further exploration to determine whether methodological and substantive features explain the differences among study findings and can account for the broad range of influences. Disparities across studies are speculated to be caused by various interventions, different analysis techniques, time spans, and geographic locations with different cultural contexts.

Table 4

Heterogeneity Analysis (Q and I^2) for Overall Mean Effect Size

	Test of heterogeneity		
	Chi-square (*Q* statistic)	*df*	Sig.
Overall	289.931	39	0.000

	Heterogeneity measures	
Overall	*Tau*-squared	0.547
	H-squared	11.456
	I-squared (%)	91.3

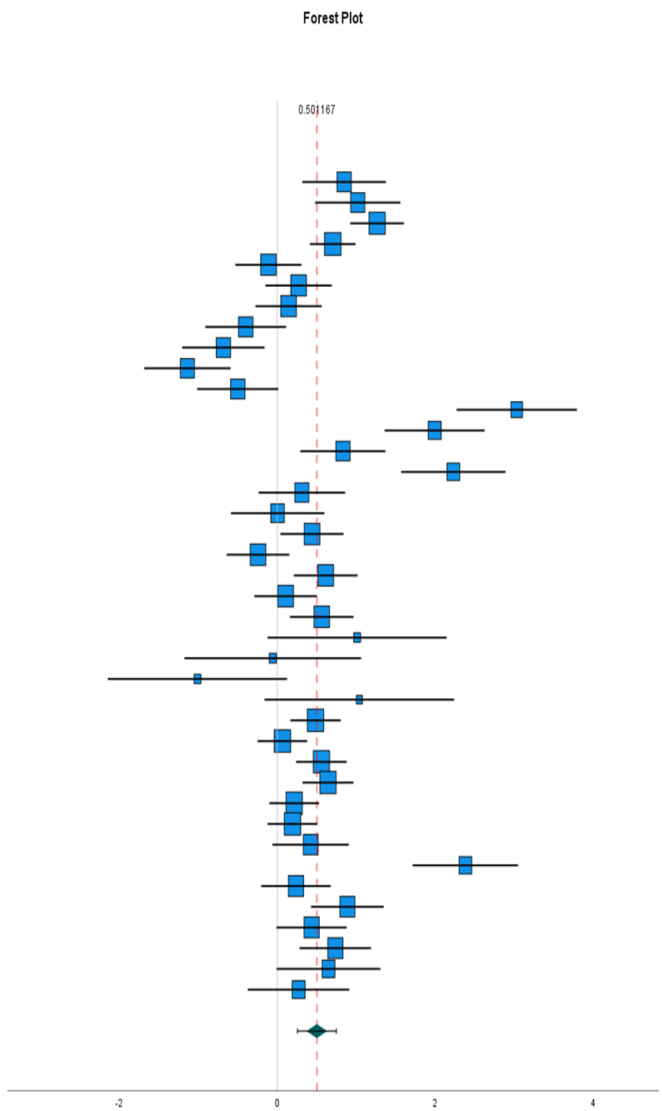
Forest Plot

The Forest plot, which illustrates the results of this meta-analysis based on the pooled forty effect sizes from the 13 included studies, is displayed in Figure 2. The study names are presented on the left of the figure. The statistics for these forty effect sizes, such as Cohens'*d*, the standard error, and the variance, are placed in the center. The Forest plot is presented on the right side of the figure. The squares centered on horizontal lines represent each study's point of estimate. Also, the squares illustrate the sizes of the studies. The size of the squares increased with the study's total number of participants. The vertical lines represent the width of the confidence intervals for each study. The confidence interval depicts the range of values likely to represent the population mean with a certain confidence level. When the line representing the sample studies in the Forest plot crosses the vertical line, the null value is contained within the 95% confidence interval, and a statistically significant difference was not observed. As reflected by the Forest plot, studies that followed a statistically significant difference are Esmaeilzad (2022), Yousef (2021), Lovelace (2016), Shavelson (2018), Vec (2019), Fischer (2009), Jimenez (2021), Petcharuk (2021), Jacob (2009), and Berry (2013). A common characteristic of these sample studies is intervention approaches used to measure critical thinking, which are the representatives of the pre and post-test. The studies that did not observe a statistically significant difference include James (2016), Hayes (2008), Fischer (2009), Jimenez (2021), Jacob (2009), and Berry (2013). These studies differed in the research design used, such as longitudinal, meaning that the sampled studies were performed over a span of time.

The studies with the broadest confidence interval are Esmaeilzad (2022) and Vec (2019). In contrast, the studies with the smallest confidence interval are Shavelson (2018), Hayes (2008), Fischer (2009), Petcharuk (2021), and Berry (2013). Thirty-two effect sizes from the included studies were classified as positive; 13% of the effect sizes revealed that the results favor interventions. The Forest plot also draws attention to a significant amount of variance when the distribution of the individual effect sizes around the average effect (the red dotted line) is not symmetrical.

Figure 2

Forest Plot of Meta-Analysis and Study-Level Statistics



Effect size of each study
Estimated overall effect size
No-effect value
Confidence interval of effect size
Overall effect size value
Estimated overall confidence interval

ID	Study	Cohen's d	Std. Error	Lower	Upper	p-value	Weight	Weight (%)
Esmaeilzad, S	UN, AN, EV	0.85	0.27	0.32	1.37	0.00	1.61	2.53
		1.02	0.27	0.48	1.56	0.00	1.61	2.52
Yousef, W.	RE,UN, AP, AN, EV	1.26	0.17	0.92	1.60	0.00	1.73	2.71
Lovelace, K	RE,UN, AP, AN, EV	0.70	0.15	0.41	0.99	0.00	1.76	2.75
James, D	UN, AN, EV, CR	-0.11	0.21	-0.53	0.31	0.60	1.69	2.64
	UN, AN, EV, CR	0.27	0.21	-0.15	0.69	0.21	1.69	2.64
	UN, AN, EV, CR	0.14	0.21	-0.28	0.56	0.50	1.69	2.64
Shavelson, R.	RE, EV	-0.40	0.26	-0.91	0.11	0.13	1.62	2.55
	EV	-0.68	0.27	-1.20	-0.16	0.01	1.62	2.54
	RE, EV	-1.14	0.28	-1.68	-0.59	0.00	1.60	2.51
	CR	-0.50	0.26	-1.01	0.01	0.06	1.62	2.54
	RE, EV	3.03	0.39	2.27	3.79	0.00	1.43	2.24
	EV	1.99	0.32	1.36	2.62	0.00	1.53	2.40
	RE, EV	0.83	0.27	0.29	1.37	0.00	1.61	2.52
	CR	2.23	0.34	1.57	2.89	0.00	1.51	2.37
Hayes, K.	AP, AN, EV	0.31	0.28	-0.24	0.86	0.27	1.60	2.51
	AP, AN, EV	0.00	0.30	-0.59	0.60	0.99	1.57	2.46
Vec, T	UN	0.44	0.20	0.04	0.84	0.03	1.70	2.66
	AN	-0.24	0.20	-0.64	0.15	0.23	1.70	2.66
	EV	0.61	0.21	0.21	1.02	0.00	1.70	2.66
	AN, EV	0.11	0.20	-0.29	0.50	0.60	1.70	2.67
	AP	0.56	0.21	0.16	0.97	0.01	1.70	2.66
Fischer, S	Memory RE	1.01	0.58	-0.12	2.14	0.08	1.13	1.78
	Memory RE,	-0.06	0.57	-1.17	1.06	0.92	1.15	1.80
	Unjustified UN,	-1.01	0.58	-2.14	0.12	0.08	1.13	1.78
	Unjustified UN,	1.04	0.61	-0.16	2.24	0.09	1.09	1.70
Jimenez J	UN, AP, AN	0.48	0.16	0.17	0.80	0.00	1.74	2.73
	UN, AP, AN	0.06	0.16	-0.25	0.38	0.69	1.75	2.74
	UN, AP, AN	0.56	0.16	0.24	0.88	0.00	1.74	2.73
	UN, AP, AN	0.64	0.16	0.32	0.96	0.00	1.74	2.73
	UN, AP, AN	0.21	0.16	-0.10	0.53	0.18	1.74	2.73
	UN, AP, AN	0.19	0.16	-0.12	0.51	0.23	1.74	2.73
Wallace, E	UN, AP, AN, EV, CR	0.42	0.25	-0.06	0.91	0.09	1.64	2.58
Petcharuk, R	UN, AN, EV, CR	2.38	0.34	1.71	3.04	0.00	1.51	2.36
Jacob, S.	RE, UN	0.24	0.22	-0.20	0.68	0.29	1.67	2.62
	AP	0.89	0.23	0.43	1.35	0.00	1.66	2.60
	AN, EV	0.43	0.23	-0.01	0.88	0.05	1.67	2.62
	CR	0.73	0.23	0.28	1.19	0.00	1.66	2.61
Berry, T.	UN, AP, AN, EV	0.65	0.34	-0.01	1.30	0.05	1.52	2.38
	UN, AP, AN, EV	0.27	0.33	-0.37	0.91	0.41	1.53	2.39
Overall		0.50	0.13	0.26	0.75	0.00		

Model: Random-effects model
Heterogeneity: Tau-squared = 0.55, H-squared = 11.46, I-squared = 0.91
Homogeneity: Q = 289.93, df = 39, p-value = 0.00
Test of overall effect size: z = 4.00, p-value = 0.00

Power Analysis

The concept of statistical power originates from traditional hypothesis testing (Bhandari, 2022b). It is related to the two types of errors that can occur in a hypothesis test. The first error is to accept the alternative hypothesis while the null hypothesis is true. This leads to a false positive, a Type I or α error. Conversely, accepting the null hypothesis is possible, while the alternative hypothesis is true. This produces a false negative, known as Type II or β error. The power of a test depends on β and is defined as Power = 1− β1. A power analysis, seen from a different angle, is a computation that aids in determining the required minimum sample size for a study.

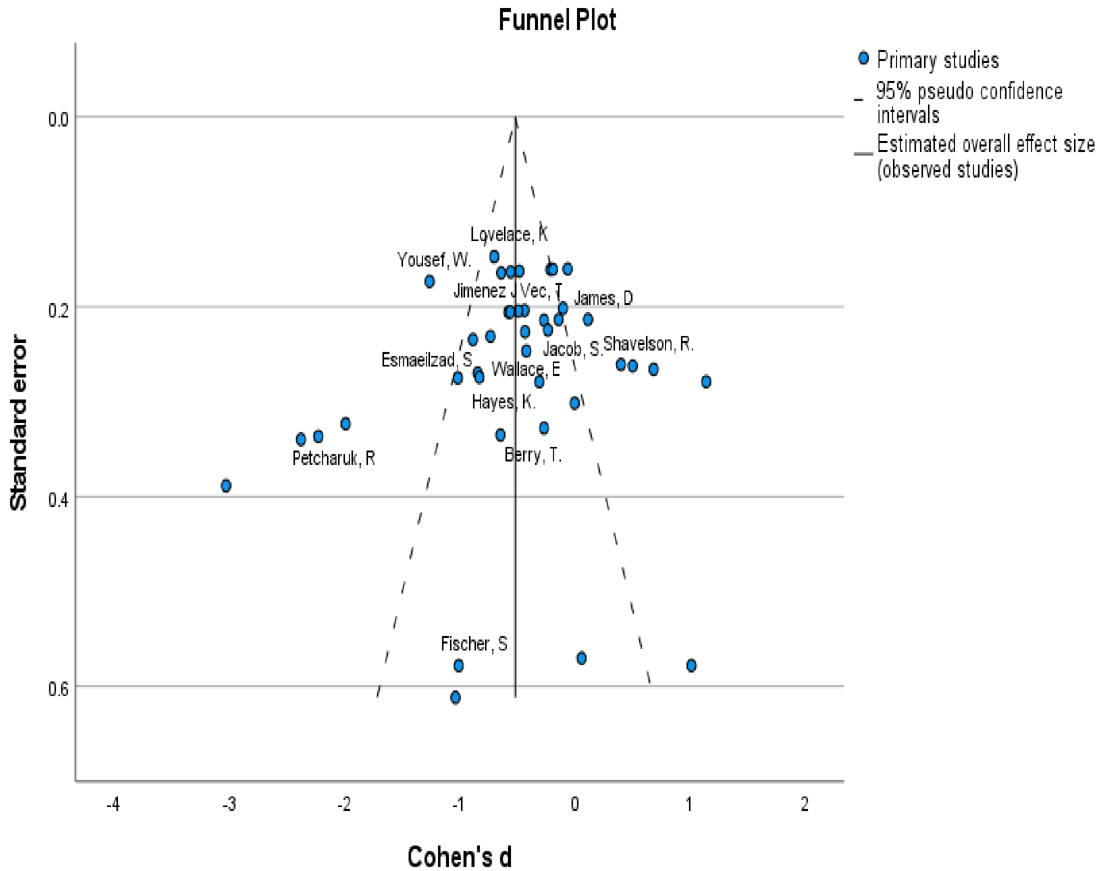
Based on the power analysis conducted using the Power Calculation for Meta-Analysis tool (Valentine et al., 2010), a sample size or studies of *N* = 9 is adequate to maintain statistical power for this random effect meta-analysis research. The actual sample size or number of effect sizes of this present study is *N* = 40. The analysis results indicated that a sample size of *N* = 9 generates adequate effect sizes to test the study hypothesis. This calculation is based on a high heterogeneity power. The number of effect sizes included in this study is greater than 9, totaling 40. The analysis indicated that the probability of committing a Type II error in the data is minimal. It is important to note that power analysis does have limitations. A common disadvantage of the power analysis is the inability to thoroughly generalize (Bruin, 2006). If the data collection method or statistical technique used to analyze the data were changed in any manner, it would be necessary to revise the power analysis to address this limitation.

Publication Bias

When studies with favorable findings are more likely to be published than those with adverse outcomes, this is known as publication bias (Higgins et al., 2022; Zhu & Carriere, 2018). Publication bias is a known pitfall in meta-analysis (However, statistical literature does not provide standardized methods to avoid such biases. This limitation made it difficult to determine an acceptable method for detecting and correcting publication bias for the study. I used the standard method, the funnel plot in IBM SPSS Statistics, to detect publication bias. The funnel plot is a scatter plot that compares the effect estimates from several studies to a measurement indicating each study's magnitude or validity (Deeks et al., 2022). The standard error of the effect estimate is used to determine the size of the study and is plotted on a vertical axis with a reversed scale to place the most extensive, significant studies at the top. The effect estimates from smaller studies will scatter more widely at the bottom, narrowing the spread among larger studies. The effect estimates from smaller studies will be more dispersed at the bottom, while the gap between larger studies will close. Without bias, the plot resembles a symmetrical (inverted) funnel. The funnel plot, as shown in Figure 3, revealed that studies are split on both sides. The inverted funnel suggests that the method used to collect data in the literature was not biased. Furthermore, it shows that bias was not introduced by overlooking small effects. The Funnel plot, presented in Figure 3, represents the data.

Figure 3

Funnel Plot



Moderating Variables

Quantitative studies frequently reveal systematic differences in effect sizes. To avoid issues with violating assumptions, researchers commonly investigate heterogeneity. Due to variations in research design, implementation of the interventions, sample size, and other factors, specific subgroups within the sampled studies may produce more significant average mean effect sizes than others (Cuijpers, 2021). This section presents a moderator analysis to examine the high variability in this analysis. Given the limited number of independent studies, the moderator analysis is described as exploratory.

Research Design

Studies with pre-and post-test experimental interventions were most prevalent in this meta-analysis. Specifically, 62% (8 of 13) of the sampled studies were pre and post-test experimental designs. The remaining 38% of sampled studies used a longitudinal, cross-sectional, or evidenced-centered design. The pre and post-test design involves measurements taken on participants under controlled conditions before and after exposure to the intervention. The studies that contained this type of design encompassed the process of selecting groups of participants, administering the critical thinking skill measurement, exposing the participants to the intervention, and re-administering and analyzing the measurement. Although the experiments were carried out under controlled conditions, control groups were not a condition of pre and post-test design in the individual studies. A drawback of not having control groups is threats to internal and external validity. Without a control group, there is a chance that the effect size will be influenced by the participant and environmental factors that are typically difficult to distinguish from the effect of the intervention (Cuijpers et al., 2017; Knapp, 2016). In some sampled studies, undetermined factors could have compromised the research design by increasing heterogeneity. Those sampled studies might have yielded a higher mean treatment effect size than those used in other designs. Despite this disadvantage of the pre and post-test design, the studies were still considered in this meta-analysis and contributed to the pool of effect sizes needed to respond to the research question adequately. The studies were based on a randomized control-group pre and post-test and followed the traditional statistical method, Analysis of variance (ANOVA). Both methods reduce the risk of threats to internal and external validity (Dimitrov & Rumrill, 2003).

Although it typically has fewer threats to internal validity than the other research designs, the between-subjects design holds limitations that could have contributed to the high variability in this analysis (Bhandari, 2022). Generally, at least one control group, one experimental group, or multiple groups vary on a characteristic in a between-subjects design (e.g., instructions, courses, training platform, and test design). An experimental group is treated with an independent variable intervention that the researcher expects will influence the outcome. In contrast, control groups receive no treatment, a standard non-related, or an imitation treatment (Maggetti et al., 2013). To determine if the manipulation of the independent variable is effective, the outcomes of the dependent variables are contrasted with the differences between groups. The researcher can conclude that the autonomous variable manipulation probably contributed to the differences if the groups diverged considerably. However, since different participants provided data for each condition, it is conceivable that the groups vary significantly between conditions and that these variations serve as alternative explanations for the findings.

Specific characteristics constrain the size and scope of cross-sectional research. Researchers usually examine distinct relationships during a particular period (Wang & Cheng, 2020). Similar to other research designs, the statistical analysis aims to evaluate the outcome for significance within a preset range. There are thus fewer threats to consider if the data begins moving in different directions. The accuracy of cross-sectional studies is constrained by the assessment techniques used by the researchers during the data collection process. For cross-sectional research to be valuable, it must be designed appropriately and representative of the complete demographic. If such a representation is not feasible, then the data collected from the participants will have inherent errors that must be considered.

Additionally, the researcher's personal bias can impact the results of cross-sectional studies. For instance, if researchers want to obtain a particular outcome, they can pose questions that will guide participants to the desired response, thereby increasing the heterogeneity level (Thomas, 2022). Also, because cross-sectional studies represent a snapshot of the condition at a single point in time, the timing of the data collection may be unrepresentative of the group's behavior, which could have been a driver of heterogeneity in the effect sizes.

A longitudinal study is the opposite of a cross-sectional study. This type of study involves repeated data collection from the same subjects over time, often focusing on a smaller group of individuals connected by a common trait. A weakness in longitudinal studies is that, over time, participants may cease to participate, impacting the sample size. In the two longitudinal studies, James (2016) retained all participants, but in Vec (2019), by the third year, eight students were unavailable at the time of the first test in the social pedagogy course. Eventually, the student received the test upon availability.

Learning and Study Strategies

Interventions that include learning, study strategies, and pre and post-tests (Lovelace, 2016; Shavelson, 2018; Petcharuk, 2021; Hayes, 2008; and Fischer, 2009) are another possible explanation for the high variation. The studies included in this meta-analysis employed a variety of instructional practices and pre and post-test instruments in a broad range of formats and applications. The nature of the variations among the instructional practices and pre and post-tests included in the sampled studies likely predisposed fluctuations in the effect sizes. High levels of variation in the sampled studies may be seen in the previously described Forest plot, shown in Figure 2. The three studies' interventions included a learning or training program and separate groups. Shavelson (2018) and Fischer employed a scenario-based training program. Conversely, Petcharuk used an interactive classroom learning environment and an information-based problem-solving learning system.

Sample Sizes

Another way effect size varies is through sample sizes. Studies with few replications will show over or under-estimations of the true effect size, leading to sampling error. The sample sizes of the sampled studies range from 19 to 215. It is inferred that the sample size variation contributed to heterogeneity within the individual studies' effect sizes.

Subgroup Analysis

A subgroup analysis was conducted to measure the differences in effect sizes between the sampled studies. The coding scheme revealed two primary measurement designs across the studies used in this analysis, pre and post-test (*n* = 10) and separate groups (*n* = 3). When comparing the effect sizes between the two types of measurements, the findings suggest that pre and post-test might be a prevalent measurement design for the higher-level features. Since the scope of this meta-analysis did not include an investigation into the effects of the types of measurement designs, future research might give more insight into which is more effective.

Analysis of Higher-Level Thinking Variables

By integrating a frequency analysis of the features related to higher-level thinking, it may be possible to explain the frequency and magnitude of the variables emphasized among the sampled studies. The variable r*emembering* appeared in 10% of the studies interventions; the variable *understanding* appeared in 22% of the included studies interventions; the variable *applying* appeared in 16% of the included studies interventions; the variable *analyzing* occurred in 22% of the included studies interventions; the variable *evaluating* occurred in 22% of the included studies interventions, and the variable *creating* appeared in 10% of the included studies interventions. The *understanding*, *analyzing*, and *evaluating* variables were the most dominant among the higher-level variables.

Summary

Results of this meta-analysis indicated that common features of higher-level thinking (remembering, understanding, applying, analyzing, evaluating, and creating) moderately contribute to developing junior and mid-level learners' critical thinking skills in postsecondary education. I^2 statistics revealed high levels of heterogeneity in terms of effect sizes between individual studies. While variations in individual studies' effect sizes can be explained, results also demonstrate the importance of consistent measures to assess critical thinking skills.

This chapter included results of this meta-analysis and a summary of findings. The study revealed a moderate overall effect size regarding the extent to which higher-level thinking skills contribute to developing critical thinking skills. In addition, tests for homogeneity and heterogeneity reinforced the use of a random effects model to obtain a true effect size. In Chapter 5, study findings are interpreted in terms of the extent to which higher-level thinking abilities promote the growth of critical thinking skills. The findings are discussed in the context of Bloom's revised taxonomy. The chapter includes an explanation of the study's limitations and further recommendations for research. The chapter will conclude by discussing the study's implications for positive social change.

Chapter 5: Discussion, Conclusions, and Recommendations

Cultivation and practice of critical thinking skills are at the forefront of 21st century intellectual requirements for the U.S. Armed Forces (Joint Chief of Staff, 2020; Bouygues, 2019; Burgard, 2020). Senior U.S. Army leadership and military experts have expressed concerns about the level of critical thinking in the armed forces, considering uncertainties and unknowns that characterize contemporary warfare (Antrobus & West, 2022). The global security environment continues to be impacted by emerging transnational threats, regional conflicts, and great power competition (JCS, 2020; Thomas, 2021). Some examples of emerging threats, regional conflicts, and power competition include Russia's invasion of Ukraine, China's attempt to seize territory throughout the South China Sea, radical terrorism, evolving air and missile threats, advanced technology in the form of drones, rockets, hypersonic weapons, acquisition of nuclear weapons, cyber disinformation, and influence campaigns and hegemonic ambitions of Russia and China (Kagan, 2016; Wright, 2021).

Another obstacle for the armed services is the current state of education, which is vital to their success and frequently connected to national security issues. The U.S. allocates more resources to education than almost every advanced country. However, students' reading performance is close to average, and their math and science performance is somewhat below average (Klein et al., 2012). Compared to their peers in Poland, Canada, New Zealand, Korea, and China, students in the U.S. are behind. The U.S. is leading toward an education crisis that will probably jeopardize the availability of public servants, particularly in specialized professions, to satisfy sociological, political, economic, and national security demands. For students to be qualified to work in institutions like the Foreign Service, intelligence community, and Armed Forces, educational institutions must enhance students' abilities and provide them with a strong basis of knowledge and cognitive skills. The State Department often experiences difficulties recruiting enough multilingual personnel (Klein et al., 2012; Srinivasan et al., 2022). U.S. generals continue to warn that enlistees cannot comprehend advanced and complicated equipment instruction manuals. Out of 250 intelligence personnel, the XVIII Airborne Corps in Iraq determined that less than five could make logical connections between ideas to form a conclusion.

To effectively navigate intricate and ever-changing operational environments, warfighters must possess a combination of professional, technical, and cognitive abilities, specifically critical thinking skills (CJCS, 2020; Odierno & McHugh, 2015; Straus et al., 2013). Although the U.S. Army PME curriculum includes critical thinking skills, U.S. Army leaders are uncertain whether the armed forces are sufficiently developing and practicing the capability of operations. The U.S. Army lacks standardized assessment tools to measure the acquisition of warfighters' critical thinking skills within the PME system.

To address this gap, I strove to examine the extent to which common features of higher-level thinking (remembering, understanding, applying, analyzing, evaluating, and creating) contribute to developing junior and mid-level learners' critical thinking skills in postsecondary education. This chapter includes key findings related to the research question. Also included is a discussion of Bloom's revised taxonomy model. The chapter concludes with a discussion of the limitations of the study, areas for future research, and a summary.

Interpretation of the Findings

Distinguishing the most salient critical thinking features might give the U.S. Army information for developing a practical critical thinking assessment approach. Critical thinking assessment strategies could assist senior leaders in comprehending warfighters' critical thinking capacities to better prepare them with the cognitive abilities required to support the military's obligation to preserve global security.

I used a meta-analysis design to collect and extract data from multiple studies based on pooling and calculating data to determine a pooled intervention effect. This meta-analysis included 13 studies. This study was based on the following research question and guided by Bloom's revised taxonomy:

RQ: To what extent can standard features of higher-level thinking (remembering, understanding, applying, analyzing, evaluating, and creating) contribute to developing junior and mid-level learners' critical thinking skills in postsecondary education?

$H_0$1: Standard features of higher-level thinking (remembering, understanding, applying, analyzing, evaluating, and creating) do not contribute to developing critical thinking skills in postsecondary education.

H_a1: Standard features of higher-level thinking (remembering, understanding, applying, analyzing, evaluating, and creating) contribute to the development of critical thinking skills in postsecondary education.

Findings in Chapter 4 substantially answered the research question.

Findings

I aimed to summarize available evidence regarding the extent to which common features of higher-level thinking develop critical thinking skills. Despite the variability of studies, the most obvious finding to emerge from this study is that common features of higher-level thinking skills (remembering, understanding, applying, analyzing, evaluating, and creating) have a moderate effect on the development of junior and mid-level learners critical thinking abilities, as supported by pooled effect sizes (Esmaeilzad, 2022; Hayes, 2008; Jimenez, 2021; Lovelace, 2016; Petcharuk, 2021; Vec, 2019; Yousef, 2021). More than half of effect sizes favor interventions with measurement approaches incorporating these features. Findings showed that high-level features *understanding*, *analyzing*, and *evaluating* emerged throughout interventions nearly half the time. The high-level feature *applying* appeared slightly less than 20% of the time. The high-level features *remembering* and *creating* occurred much less than *applying*. This suggests learners might have more difficulty developing and demonstrating these skills within timeframes and conditions of interventions. Students need time, education, and expertise to build mental structures of higher-level thinking. A study on the length of exposure and estimated duration to build cognitive abilities between the higher-level thinking levels can be a future research topic.

Findings reinforced the connection between Bloom's revised taxonomy model and assessing learners’ critical thinking ability. The sampled studies analyzed in this research provided credence that Bloom's revised taxonomy model is a viable framework for instructors to measure students' critical thinking abilities through formative and summative assessments (Farcis et al., 2022; Zapalska et al., 2018). Since developing and advancing critical thinking skills cannot be accomplished without some instructional and evaluation support, Bloom's revised taxonomy model functions as a tool to create teaching and development approaches. Furthermore, Bloom's revised taxonomy is used to help create a balance between assessment and evaluation activities to ensure that students' learning involves practicing higher-level thinking. This study supports the view that learners may improve their critical thinking abilities after controlling for other variables, guided by the common features of higher-level thinking.

As expected, the studies varied in the types of measurements used to assess critical thinking skills, which are likely attributed to the observed heterogeneity. The measurements included questionnaires, tests or examinations, open-ended essay questions, multiple choice questions (MCQ), and project or lecture activity observations. The MCQ format is the least favored, whereas the open-ended essay question format is the ideal measurement (Dwyer, 2018; Goodsett, 2020). Moderators claim that the MCQ is the least preferred since it enables test-takers to guess the proper response rather than proving their capacity to critically examine, assess, and deduce solutions to issues (Goodsett, 2020; Ku, 2009). The open-ended questions allow test-takers to exhibit their ability to employ critical thinking skills instantaneously.

Although not within the scope of this research, an examination of the types of measurement instruments that better align with assessing the higher-level features revealed that pre and post-tests and separate groups are most prevalent across the studies. The subgroup analysis indicated that the pre and post-test weighed the heaviest and is consistent with other studies that showed successful interventions in developing critical thinking. According to the subgroup analysis, the pre-test/post-test using the common features carried the most weight, which aligns with previous research in the literature showing the effectiveness of interventions in promoting critical thinking.

Shavelson (2019) suggested that critical thinking assessment needs to be improved in the literature. According to the author, concerns about the quality of student performance in critical thinking, problem-solving, perspective-taking, and communication, referred to as twenty-first-century skills, have arisen due to the rapid proliferation and broad range of higher education. (Later studies revealed that the perception of the need for critical thinking assessments is more than a local matter (Al-Mahrooqi & Denman, 2020; Goodsett, 2020; Mohd et al., 2019). The researchers asserted that while developing students' critical thinking abilities is frequently seen as a key goal of higher education, policymakers, educators, and students in some universities in the Middle East and North Africa region have expressed concern about this issue (Al-Mahrooqi & Denman, 2020). One challenge is that little research has been performed on measuring learners’ critical thinking skills.

Limitations of the Study

This study has various strengths, which were discussed in the findings section. However, due to the unconventional research design and the scope of the study, limitations need to be considered. The meta-analysis design provided a transparent and structured approach to combine and analyze data from different studies and present results. Contrarily, it was challenging to include studies because of the requirements for the meta-analysis inclusion criteria. The literature selection approach used in this meta-analysis is partially attributable to the small sample size. During the initial steps of the data collection process, more than 122 primary studies from the literature search were excluded due to insufficient quantitative data to estimate effect size. This meta-analysis only includes forty total effect sizes, which is inadequate to thoroughly investigate whether potential effect modifiers explain heterogeneity between the studies. Meta-analytical results must be interpreted considering the structure and level of detail provided in the primary studies.

While the primary studies provided sufficient data to analyze the effect sizes, they varied in the narratives of factors that might have influenced the study. Although the inclusion and exclusion criteria aid in determining the quality of the studies, the researchers' drive and analytical and interpretive abilities impact data reporting (Higgins, 2022). As a final point, the selected variables for insight were limited to the elements of Bloom's revised taxonomy model. An additional investigation is worthwhile to discover if a broad assessment strategy that includes these common features can be developed. Furthermore, variables such as gender, age, location, or educational attainment were not included but could have provided additional information on how these affected the findings.

Recommendations

Returning to the research question, the present study's findings authenticate a set of core higher-level thinking features for the U.S. Army to formulate generalized assessment or measurement strategies to gauge the acquisition and practice of warfighters' critical thinking skills. The following recommendation is offered from the perspective of the research, analysis, and findings presented throughout this study.

Recommendation 1: Instruments to Measure Critical Thinking

A natural progression of this study is to research the practicality of creating a structured generalized model to measure warfighters' critical thinking skills by integrating these core higher-level thinking features. The value of developing assessment tools that incorporate these fundamental aspects of higher-level thinking will become apparent in the effectiveness of measuring critical thinking proficiency within the U.S. Army formations and in the ability to shape training or instructional requirements. A benefit of using these core features is that they are flexible and adaptive for assessing performance based on learners' levels of achievement and diverse needs (Sideeg, 2016). Explicit, written criteria will help consistently evaluate materials and provide transparency and clarity throughout the process.

Using this recommendation, an example of an assessment model that might be created to integrate higher-level thinking features is an effective algorithm. The algorithm should outline the expectations for each level within the taxonomy. This ensures that the instructor and learner understand the performance indicators for each assessment activity. The algorithm should at least include the criteria, descriptors, and scoring. The criteria define the key elements that comprise each of Bloom's revised taxonomy levels. The descriptors describe what is required at each level of performance with the assessment activity. Scoring means assigning a numerical value or range for each performance descriptor, facilitating consistent and objective grading. The scoring can be represented through letter grades, number of points, or descriptors of quality levels, such as exemplary, competent, and developing. It is important to note that any assessment must align with the instruction, learning activities, and learning outcomes (Biggs & Tang, 2011; Maki, 2010).

Recommendation 2: Add Critical Thinking Skills Assessment to Policy

Policy plays a crucial role in forming and implementing guidelines and principles of an institution. Policies exert a powerful influence. Institutional goals may be accomplished with the support of policy implementation. With it, goals and objectives of foremost importance might culminate in real or sufficient change at the operational levels. Therefore, broadening current policy, which requires the PME facilitators to inculcate critical thinking skills into its curriculum, incorporating measurement approaches, will help assess this vital ability within U.S. Army formations. A revived policy that promotes investment at all levels of creating a force with a critical-thinking mindset might also assist the goal of an adaptable, formidable force. A force capable of dealing with complexity, ambiguity, uncertainty, and rapid change while still being able to make tough decisions under pressure.

Recommendation 3: Invest in Critical Thinking Measurement Instruments

Funding the development of assessment systems that employ various instruments to gauge the critical thinking abilities of warfighters should be a priority for U.S. Army policymakers. Adopting assessment techniques that are thorough and successfully connect with course objectives and instructions is vital to acquire valuable data on warfighters' critical thinking skills. The quality of thinking abilities cannot be immediately judged after receiving two to three repetitions of instructions. Holmes et al. (2015) claim that evaluating critical thinking is a process rather than a quick exercise. The purpose of assessments is to provide instructors, school administrators, and students with discrete formative data on difficult-to-measure skills.

Furthermore, given the emphasis on developing critical thinking skills and its impact on the U.S. Army’s effectiveness in exerting land power, policymakers must implement a policy that promotes assessment systems in the operational training and PME. The assessment systems need to focus on measuring the growth or progress of students and instructions over time. A policy can also help curriculum developers, training facilitators and instructors, and other stakeholders understand the importance of facilitating training and learning environments to inculcate the learning and practice of critical thinking. The U.S. Army leaders can apply this study as a roadmap to orient policy toward establishing and integrating measurement instruments to demonstrate the application of cognitive abilities and, importantly, critical thinking skills across training and education programs and in the evaluation reports.

Recommendation 4: Improved Practical Training for PME Instructors

Businesses have often urged educators to improve how they teach students to think critically. In a report published in 2019 by the Society for Human Resource Management, 51% of businesses surveyed claimed that educational institutions had done little to close the soft skills gap (Society for Human Resource Management. 2019). As a professional institution, part of the military's function is to educate its forces. The future of the military profession depends on fostering and advancing knowledge and thinking skills, much like businesses. It takes time and effort to develop critical thinking skills, and it demands educators who can modify the learning environment to promote intellectual and cognitive development through rigorous and timely education and training. Instructors must thus be educated and trained to promote critical thinking skills by encouraging in-depth discourse, debate, and discussion. Instructors must be conscious of their thought patterns, practice critical thinking, and model the necessary behaviors that emerge from training and experience (Holmes et al., 2015). It is unrealistic to anticipate that merely being exposed to a subject will teach students how to think critically. Students require proper teaching from experienced professionals if they are to be expected to develop the skill. To effectively teach students and execute formative assessments in the classroom, instructors have a responsibility to grow a variety of professional competencies. This may be achieved by receiving practical instruction and continuing to advance competencies.

Recommendation 5: Future Research on Multiple Generational Learning

The military will benefit from future research that broadens the scope of this study by accounting for generation differences and learning preferences in educational and training programs that evolve cognitive development. Even though it is commonplace that military personnel originate from different generations, the environment and life experiences of the individual influence their learning and development differently. For example, the first generations to use digital technology daily are Millennials and Generation X (Sutton, 2019). Members of these generations are known to favor interactive, dynamic learning and want social media to be interwoven into their learning environment. If learning models and teaching techniques do not apply to learners or are sufficiently connected to their socio-educational realities, they lose their value (López-Noguero, 2008; Marín-Díaz, 2015). The military, notably the U.S. Army, must consider the generational disparities and customize curricula, training programs, and assessment methods to accommodate learners' learning preferences and styles.

Also, it would be interesting to explore whether learning models such as gamification and classroom escape rooms appeal to the diverse needs of learners. Research of this nature might also provide insight into improving educational and training practices and assessments to accommodate multigenerational learning in the military. Gamification refers to using game mechanics to make learning more engaging for students (Buljan, 2021). Gamification fosters the development of abilities, including problem-solving, critical thinking, social awareness, cooperation, and teamwork. Further, engaging in games can boost students' motivation, arouse their interest in certain subjects, and enhance their cognitive function. Classroom escape rooms are real-time cooperative games where players must overcome challenges to fulfill objectives under predetermined time constraints (Velcamp et al., 2020). Additionally, these educators pointed out that escape rooms provide an environment for individuals to develop or improve their collaboration, communication, problem-solving, critical thinking, and analytical reasoning abilities. These learning models are continuously emerging as innovative approaches to teaching, training, and assessing the acquisition of students' skills.

Implications

The results of this meta-analysis research have significant implications for how the U.S. Army and the military can better equip warfighters with the cognitive capabilities, primarily critical thinking skills, to function in the future security environment.

Implications for Future Practice and Research

Critical thinking skills are among the most demanded competencies of the 21st century for the military and society. This research addresses U.S. Army leaders' uncertainty about the quality or performance of warfighters' employment of critical thinking. Compared to other studies, this research used a thorough and methodologically sound procedure to combine the most reliable, quality data and produce evidence of core features to measure or assess critical thinking skills.

The implications of these findings are intended to improve the cognitive capabilities, specifically critical thinking, of the U.S. armed forces and the military. The validation of the common features of higher-level thinking skills: remembering, understanding, applying, analyzing, evaluating, and creating can be used as a roadmap to an effective critical thinking assessment strategy that will better equip military personnel with cognitive tools to adapt to complex situations. This perspective can enhance the Army's ability to deliver and sustain an unmatched scale of ground forces to advance the protection of our nation.

Social Change

The current study could contribute to social change by helping educators and those invested in the reform to shift education systems to equip learners with 21st-century competencies such as critical thinking. It is reasonable to expect that the U.S. Army PMES facilitators and the military will use these core features to create constructive assessment strategies that adequately gauge warfighters' critical thinking skills. Following a competency standard can help improve the PMES program. With a standard of this kind in place, facilitators can modify the learning objectives and training requirements specific to the learning environment and implement best practices that aim to improve critical thinking competencies.

Conclusion

The responsibility for securing a broad range of demanding national security and defense objectives falls upon the U.S. Armed Forces. Equally, the Army forces will continue to face a wide range of threats, from near-peer contests to potential encounters with violent extremist organizations. The combination of these global developments signals that Army forces of today and the future will have more cognitive demands, pointedly critical thinking, placed upon them. These demands involve the capacity to reason, prioritize competing initiatives, identify novel solutions, select the optimal risk-based choices, and swiftly adjust to shifting circumstances while conducting mission objectives. An essential premise for efforts to achieve the Joint Chiefs of Staff and Senior Military leadership's vision of a critical-thinking joint warfighting force is for facilitators of the Professional Military Education System to develop a curriculum to train and develop these skills.

This meta-analysis explored the extent to which the common features of higher-level thinking, remembering, understanding, applying, analyzing, evaluating, and creating, contribute to developing junior and mid-level learners' critical thinking skills in post-secondary education. The importance of validating this hypothesis is to demonstrate a baseline set of common features or indicators that the PMES can use to create a measurement approach to validate warfighters' critical thinking skills and capabilities. The structured data collection and detailed analysis yielded the result that, in general, the common higher-level features moderately lead to developing learners' critical thinking skills. This study substantiates that developing junior and mid-level learners' critical thinking through instructional or training interventions augmented by measurement approaches is possible.

Cognitive capabilities, particularly critical thinking, do not develop naturally. Still, they must be fostered gradually through specifically developed processes and measurement instruments composed of standardized core features to appropriately assess the skills learned and transfer them into operational environments. Given the current and future security environment characterized as complex, unpredictable, and dynamic, the U.S. Armed Forces must rely on more than sophisticated weapons systems and technology. The increased demands upon our forces require leaders at all levels who think critically. This study offers the best representation of a set of core common higher-level thinking features that can lead to practical evaluation assessment approaches to validate critical thinking capability throughout the U.S. Army formations. This study reinforces the necessity for the military stakeholders and those in the educational and social science fields to invest in the pursuit of reliable and valid assessment strategies using the core higher-level thinking skills for measuring warfighters' thinking skills.

References

Abrami, P. C., Bernard, R. M., Borokhovski, E., Wade, A., Surkes, M. A., Tamim, R., & Zhang, D. (2008). Instructional interventions affecting critical thinking skills and dispositions: A stage 1 meta-analysis. *Review of Educational Research. 78*(4). pp. 1102-1134. https://doi.org/10.3102/0034654308326084

Abrami, P. C., Bernard, R. M., Borokhovski, E., Waddington, D. I., Wade, C. A., & Persson, T. (2015). Strategies for teaching students to think critically: A meta-analysis. *Review of Educational Research. 85*(2). pp.275–314. https://doi.org/10.3102/0034654314551063.

Abosalem, Y. (2016). Assessment techniques and students' higher-order thinking skills. *International Journal of Secondary Education. 4*(1). Pp. 1-11. https://doi.org/10.11648/j.ijsedu.20160401.11.

American Management Association (2019). AMA critical skills survey: Workers need higher level skills to succeed in the 21st century. https://www.amanet.org/articles/ama-critical-skills-survey-workers-need-higher-level-skills-to-succeed-in-the-21st-century/.

Aubteen Darabi, A. & Arrington, T.L. (2017). *Designing instruction for critical thinking: A case of a graduate course on evaluation of training*. (EJ1150794). https://files.eric.ed.gov/fulltext/EJ1150794.

Adams, M. H., Whitlow, J. F., Stover, L. M., and Johnson, K. W. (1996). Critical thinking as an educational outcome: An evaluation of current tools of measurement. *Nurse Educator. 21*(3). pp. 23–32.

https://doi.org/10.1097/00006223-199605000-00009.

Advance Consulting for Education. (2013). *Teaching critical thinking skills*. [Video]. https://www.bing.com/videos/search?q=Critical+Thinking+Techniques&&view=detail&mid=5909646EA5B3FEA65FD55909646EA5B3FEA65FD5&&FORM=VRDGAR&ru=%2Fvideos%2Fsearch%3Fq%3DCritical%2BThinking%2BTechniques%26FORM%3DVDMHRS

Al-Ghadouni, A. B. N. (2021). Instructional approaches to critical thinking: An overview of reviews. *Revista Argentina de Clinica Psicologica. 30.* pp. 240-246. https://doi.org/10.24205/03276716.2020.2020

Allen, C. D., & Gerras, S. J. (2009). Developing creative and critical thinking. *Military Review*. November-December. pp. 77-83. https://apps.dtic.mil/sti/pdfs/ADA515842

Allrich, R. (2017). Meaningful learning. https://web.ics.purdue.edu/~rallrich/learn/mean.html.

Al-Mahrooqi, R., & Denman, C. J. (2020). Assessing students’ critical thinking skills in the humanities and sciences colleges of a Middle Eastern university. *International Journal of Instruction, 13*(1), 783-796. https://doi.org/10.29333/iji.2020.13150a

Anderson, L. W., Krathwohl, D. R., & Bloom, B.S. (Eds.). (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom’s taxonomy of educational objectives*. Allyn & Bacon.

Angelo, T. A. (1995). Classroom assessment for critical thinking. *Teaching of Psychology, 22*(1), 6–7. https://doi.org/10.1207/s15328023top2201_1

Antrobus, S. & West, H. (2022). 'This is all very academic': Critical thinking in professional military education. *RUSI Journal*. *167*(3). pp. 78-86. https://doi.org/10.1080/03071847.2022.2112521

Austin, A. (2021). Secretary of Defense remarks for the U.S. INDOPACOM change of command. https://www.defense.gov/News/Speeches/Speech/Article/2592093/secretary-of-defense-remarks-for-the-us-indopacom-change-of-command/

Ayers, R. B. (2016) Optimizing workforce performance: Perceived differences of army officer critical thinking talent across the level of education. 329. https://aquila.usm.edu/dissertations/329

Bacon, F. (1605). The advancement of learning. (2016). In W. Engel, R. Loughnane, & G. Williams (Eds.), *The memory arts in Renaissance England: A critical anthology* (pp. 197-201). Cambridge University Press.

Bailin, S., Case, R., & Coombs, J. R. (1999). Common misconceptions of critical thinking. *Journal of Curriculum Studies. 31*(3). pp. 269–283. https://doi.org/10.1080/002202799183124

Barry, B. (2017). *Harsh Lessons: Iraq, Afghanistan, and the Changing Character of War. The International Institute for Strategic Studies*. [Video]. YouTube. https://www.youtube.com/watch?v=yfwiFYTq7YE.

Barry, C. L., Horst, S. J., Finney, S. J., Brown, A. R., & Kopp, J. P. (2010). Do examinees have similar test-taking effort? a high-stakes question for low-stakes testing. *International Journal of Testing. 10.* pp. 342–363. https://doi.org/10.1080/15305058.2010.508569.

Basu, A. (2014). *Introduction to Meta-Analysis*. 10.7287/peerj.preprints.665v1.

Bataineh, O., & Alazzi, K. (2009). *Perceptions of Jordanian secondary schools' teachers Towards critical thinking. Education Resources Information Center*. (EJ869429). ERIC?

Becker, B. J. (2000). Multivariate meta-analysis. In H. E. A. Tinsley & S. D. Brown (Eds.), *Handbook of applied multivariate statistics and mathematical modeling* (pp. 499–525). Academic Press. https://doi.org/10.1016/B978-012691360-6/50018-5

Becker, J.J. & DeFoor, J. E. (2018). *Exploring the future operating environment. Joint Force Quarterly 89* (2nd Quarter 2018). https://ndupress.ndu.edu/Portals/68/Documents/jfq/jfq-89/jfq-89_121-125_Becker-DeFoor.pdf?ver=2018-04-11-125441-307

Bensley, D.A., Rainey, C., Murtagh, M.P., Flinn, J.A., Maschiocchi, C., Bernhardt, P.C., & Kuehne, S. (2016). Closing the assessment loop on critical thinking: The challenge of multidimensional testing and low test-taking motivation. *Thinking skills and creativity. 21*. 158-168. https://doi.org/10.1016/j.tsc.2016.06.006.

Bellis, M. (2020). Benjamin Bloom: Critical Thinking and Critical Thinking Models. ThoughtCo. thoughtco.com/benjamin-bloom-critical-thinking-models-4078021

Benjamin, R. & Chun, M. (2003). A New Field of Dreams: The Collegiate Learning Assessment Project. The Peer Review. 5(4). https://www.aacu.org/publications-research/periodicals/new-field-dreams-collegiate-learning-assessment-project.

Berry, T. M. (2013). *Using a cognitive approach to improving critical thinking in LPN students* (Order No. 3603526). Available from ProQuest One Academic. (1469609787). https://www.proquest.com/dissertations-theses/using-cognitive-approach-improving-critical/docview/1469609787/se-2.

Bhandari, P. (2022a). Between-Subjects Design | Examples, Pros & Cons. *Scribbr*. https://www.scribbr.com/methodology/between-subjects-design/.

Bhandari, P. (2022b). Statistical Power and Why It Matters: A Simple Introduction. *Scribbr*. https://www.scribbr.com/statistics/statistical-power/.

Bhandari, P. (2023). Descriptive Statistics | Definitions, Types, Examples. *Scribbr*. https://www.scribbr.com/statistics/descriptive-statistics/.

Biddle, T. D. (2016). Making sense of the "long wars" --advice to the U.S. Army. *Parameters*, *46*(1), 7+. https://auth.waldenulibrary.org/ezpws.exe?url=http://go.galegroup.com.ezp.waldenulibrary.org/ps/i.do?p=EAIM&sw=w&u=minn4020&v=2.1&it=r&id=GALE%7CA458916150&asid=11af6105f45c077506fab2bf08ae501b

Bloom, B. S. (Ed.). (1956). *Taxonomy of educational objectives: The classification of educational goals, by a committee of college and university examiners*. New York: Longmans.

Bloom, B. (1974). *The Taxonomy of Educational Objectives: Affective and Cognitive Domains*. David McKay Company, Inc.

Bobko, P., & Stone-Romero, E. F. (1998). Meta-analysis may be another useful research tool, but it is not a panacea. In G. R. Ferris (Ed.), *Research in personnel and* *human resources management, Vol. 16*, 359–397. Elsevier Science/JAI Press.

Bokhove, C. & Campbell, R. (2020). Revising opinions about Bloom's taxonomy. *Impact Journal of the Chartered College of Teaching.* https://impact.chartered.college/article/revising-opinions-about-blooms-taxonomy/.

Bonwell, C. & Eison, J. (1991). Active Learning: Creating Excitement in the Classroom. Information Analyses - ERIC Clearinghouse Products (071). 3. ISBN 978-1-878380-08-1. ISSN 0884-0040.

Borenstein, M., Hedges, L. V., Higgins, J. P., & Rothstein, H. R. (2021). *Introduction to meta-analysis*. John Wiley & Sons.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. John Wiley & Sons Ltd.

Bouygues, H.L. (2019). U.S. Military Leaders Want Soldiers To Think Critically, Not Just Follow Orders. *Forbes*. https://www.forbes.com/sites/helenleebouygues/2019/01/10/u-s-military-leaders-want-soldiers-to-think-critically-not-just-follow-orders/?sh=5a16930ef595

Bralley, N.H. (2006). ILE: A new system for CGSC students. *Army Logistician*. 38 (1). https: www-cgsc.army.mil/carl/download/csipubs/infantry/inf_intro_cvii.

Brown, R. B. (2014). *The Human Dimension: A Framework for Optimizing Human Performance* [White Paper]. The United States Army Combined Arms Center, Fort Leavenworth, KS: Department of the Army. https://usacac.army.mil/sites/default/files/publications/Human%20Dimension%20

White%20Paper%20%28Combined%20Arms%20Center%2009%20Oct%2014%29_1.

Bruin, J. (2006). New test: command to compute new test. *UCLA: Statistical Consulting Group*. https://stats.oarc.ucla.edu/stata/ado/analysis/.

Buljan, M. (2021). Gamification For Learning: Strategies and Examples. *eLearning Industries*. https://elearningindustry.com/gamification-for-learning-strategies-and-examples.

Burgard, A.S. (2020). Applying Bloom's Taxonomy to Coast Guard Talent Management. *U.S. Naval Institute*. *Vol. 146*:9/1,411. Applying Bloom's Taxonomy to Coast Guard Talent Management | Proceedings - September 2020 *146*/9/1,411 (usni.org).

Caine, B. T. (n.d.). *Military Professional Education System*. https://education.stateuniversity.com/pages/2237/Military-Professional-Education-System.

Cañas, J. (2006). *Cognitive Flexibility*. 10.13140/2.1.4439.6326. https://www.researchgate.net/publication/272022148_Cognitive_Flexibility

Card, N. A., & Casper, D. M. (2013). Meta-Analysis and Quantitative Research Synthesis, in Todd D. Little (ed.). *The Oxford Handbook of Quantitative Methods in Psychology: Vol. 2: Statistical Analysis*, *Oxford Library of Psychology* (2013; online ed, Oxford Academic, 1 Oct. 2013). https://doi.org/10.1093/oxfordhb/9780199934898.013.0030

Care, E., Kim, H., Vista, A. & Anderson, K. (2019). Education system alignment for 21st-century skills: Focus on Assessment. *Brookings Institute*. https://www.brookings.edu/research/education-system-alignment-for-21st-century-skills/.

Case, R. (2013). Unfortunate Consequences of Bloom's Taxonomy. Social Education. *National Council for the Social Studies. 77*(4). https://www.socialstudies.org/social-education/77/4/unfortunate-consequences-blooms-taxonomy.

Castle A. (2003). Demonstrating critical evaluation skills using Bloom's taxonomy. *British Journal of Therapy and Rehabilitation*. *110.* pp.369-373. 10.12968/bjtr.2003.10.8.13515.

Chairman, Joint Chiefs of Staff. (2005) Officer Professional Military Education Policy. *Chairman of the Joint Chiefs of Staff Instruction 1800.01C*. Joint Staff, J-7, Joint Education Branch. https://www.jcs.mil/Library/CJCS-Instructions/.

Chairman, Joint Chiefs of Staff (2006). Joint Vision 2020. *National Defense University, Institute for National Strategic Studies*. https://www.jcs.mil/Portals/36/Documents/Doctrine/education/jcs_pme_tm_vision.pdf?ver=2020-05-15-102429-817.

Chairman of the Joint Chiefs of Staff. (2009). *Officer Professional Military Education Policy* [Instruction]. https://usacac.army.mil/sites/default/files/documents/cace/LREC/2011_CJCSI_1800.01D_Ch1_OPMEP.

Chairman of the Joint Chiefs of Staff. (2009). *Officer Professional Military Education Policy (OPMEP) (CJCSI 1800.01D)*. https://usacac.army.mil/sites/default/files/documents/cace/LREC/2011_CJCSI_1800.01D_Ch1_OPMEP.

Chairman of the Joint Chiefs of Staff. (2015). *Officer professional military education policy* [Instruction]. Chairman of the Joint Chiefs of Staff Instruction 1800.01E. Department of Defense. https://www.jcs.mil/Portals/36/Documents/Doctrine/education/cjcsi1800_01e.pdf?ver=2017-12-29-142206-877.

Chairman, Joint Chiefs of Staff. (2016). *Joint Concept for Human Aspects of Military Operations (JC-HAMO)*. Washington, DC: U.S. https://nsiteam.com/social/wp-content/uploads/2017/01/20161019-Joint-Concept-for-Human-Aspects-of-Military-Operations-Signed-by-VCJCS.

Chairman of the Joint Chiefs of Staff. (2020). 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*. https://www.jcs.mil/Portals/36/Documents/Doctrine/education/jcs_pme_tm_vision.

Chen, Z., & Lawson, R. B. (1996). Groupthink: Deciding with the leader and the devil. *Psychological Record, 46*(4). p. 581. https://www.thefreelibrary.com/Groupthink%3a+deciding+with+the+leader+and+the+devil.

Chiarini, A, Found, P., & Rich, N. (2015). Understanding the Lean Enterprise: Strategies, Methodologies, and Principles for a More Responsive Organization. *Cham: Springer. 132*. ISBN 978-3-319-19994-8

Christy, J., Sami, A. and Arumugam, U. (2020). A Descriptive Analysis of Students Learning Skills Using Bloom's Revised Taxonomy. *Journal of Computer Science 2020, 16* (2): 183.193 DOI: 10.3844/jcssp.2020.183.193

Cohen J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NY: Lawrence Erlbaum.

Cohen, M.S., Salas, E. & Reigal, S. (2002). *Critical Thinking: Challenges, Possibilities, and Purpose*. Cognitive Technologies, Inc. https://www.researchgate.net/publication/236030685_CRITICAL_THINKING_CHALLENGES_POSSIBILITIES_AND_PURPOSE.

Cole, J. S., & Osterlind, S. J. (2008). Investigating differences between low- and high-stakes test performance on a general education exam. *The Journal of General Education. 57.* pp. 119–130. https://doi.org/10.2307/27798099.

Cojocar, W. J. (2011). Adaptive Leadership in the Military Decision-Making Process. *Military Review*. Nov-Dec. https://www.armyupress.army.mil/Portals/7/military-review/Archives/English/MilitaryReview_20120630MC_art007.

Cooper, J. L. (1995). Cooperative learning and critical thinking. *Teaching of Psychology, 22*(1). pp. 7-8. https://doi.org/10.1207/s15328023top2201_2.

Cooper, J. M. & Hutchinson, D. S. (Eds). (1997). *Plato*. Plato: Complete Works. Hackett

Cuijpers, P., Weitz, E., Cristea, I., & Twisk, J. (2017). Pre-post effect sizes should be avoided in meta-analyses. *Epidemiology and Psychiatric Sciences. 26*(4). pp. 364-368. doi:10.1017/S2045796016000809.

Cuijpers, P., Griffin, J. W., & Furukawa, T. A. (2021). The lack of statistical power of subgroup analyses in meta-analyses: a cautionary note. *Epidemiology and psychiatric sciences. 30(*78). https://doi.org/10.1017/S2045796021000664.

Davitch, J. M., & Folker Jr., R. D. (2017). Operationalizing Air Force Critical Thinking. *Air & Space Power Journal. 31*(4). pp 62–68. https://www.airuniversity.af.edu/Portals/10/ASPJ/journals/Volume-31_Issue-4/V-Davitch_Folker.

Deeks, J.J., Higgins, J.P.T, Altman, D.G. (editors). (2022). Chapter 10: Analyzing data and undertaking meta-analyses. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). *Cochrane Handbook for Systematic Reviews of Interventions.* Version 6.3. https://training.cochrane.org/handbook/current/chapter-10.

Department of the Army. (2015). The United States Army Learning Concept for 2015. Professional Military Education. *Training and Doctrine Command Pamphlet 525-8-2*. Washington, DC: U.S. Government Printing Office. https://adminpubs.tradoc.army.mil/pamphlets/TP525-8-2.

Department of the Army, Training and Doctrine Command (2008). U.S. Army Study of the Human Dimension in the Future 2015-2024. *Training and Doctrine Command Pamphlet 525-3-7-01*. Fort Monroe, VA: Army Capabilities Integration Center.

https://apps.dtic.mil/sti/citations/ADA489116

Department of the Army, Training and Doctrine Command (2014). The U.S. Army Human Dimension in the Future 2015-2024. *Training and Doctrine Command Pamphlet 525-3-7*. Fort Monroe, VA: Army Capabilities Integration Center. https://usacac.army.mil/sites/default/files/publications/Human%20Dimension%20White%20Paper%20%28Combined%20Arms%20Center%2009%20Oct%2014%29_1.

Department of the Army, Training and Doctrine Command (TRADOC) (2018a). Army Learning, Army Education Processes. *Department of the Army Headquarters, United States Army, TRADOC Pamphlet 350-70-7*. https://adminpubs.tradoc.army.mil/regulations/TR350-70.

Department of the Army Training and Doctrine Command. (2018b). The U.S. Army in Multi-Domain Operations 2028. *Department of the Army Headquarters, TRADOC Pamphlet 525-3-1*. https://adminpubs.tradoc.army.mil/pamphlets/TP525-3-1

Department of the Army. (2018c). Officer Training and Education. Army Training and Leader Development. *Army Regulation 350-1*. https://armypubs.army.mil/epubs/dr_pubs/dr_a/pdf/web/arn18487_r350_1_admin_final.

Department of the Army. (2019). The operational environment and the changing character of warfare. *Department of the Army Headquarters, TRADOC PAM 525-92*. https://adminpubs.tradoc.army.mil/pamphlets/TP525-92.

Devedzic, V., Tomic, B., Jovanovic, J., Kelly, M., Milikic, N., Dimitrijevic, S., Djuric, D., & Sevarac, Z. (2018). Metrics for Students' Soft Skills. *Applied Measurement in Education. 31*(4). pp. 283–296. https://doi-org.ezp.waldenulibrary.org/10.1080/08957347.2018.1495212.

Dempsey, M. (2012). Chairman Joint Chiefs of Staff, Joint Education [White Paper]. http://www.jcs.mil/content/files/2012-07/071812110954_CJCS_Joint_Education_White_Paper.

Dempsey, M. E. (2012). Mission Command [White Paper]. www.jcs.mil/content/files/2012-04/042312114128_CJCS_Mission_Command_White_Paper_2012_a.

Dewey, J. (1910). *How We Think.* New York: D. C. Heath and Company. doi:10.1037/10903-000. http://books.google.com/books?id=WF0AAAAAMAAJ&pg=PP1

Dimitrov, D. M., & Rumrill, P. D., Jr (2003). Pretest-posttest designs and measurement of change. *20*(2). pp. 159–165. https://www.researchgate.net/publication/10826237_Pretest-Posttest_Designs_and_Measurement_of_Change

Dwyer, C.P., Hogan, M.J. & Stewart, I. (2014). An integrated critical thinking framework for the 21st century. *Thinking Skills & Creativit*y. *12.* pp. 43-52. https://www.sciencedirect.com/science/article/abs/pii/S1871187114000030

Dwyer, C.P. (2017). *Critical thinking: Conceptual perspectives and practical guidelines*. Cambridge University Press.

Egan, B.D. (2005). *The role of critical thinking in effective decision mak*ing. Global Knowledge. https://articulosbm.files.wordpress.com/2010/03/criticalthinking

Elder, L. (2007). Our Concept and Definition of Critical Thinking. *The Foundation for Critical Thinking*. http://www.criticalthinking.org/pages/defining-critical-thinking/766.

Elder, L. & Cosgrove, R. (n.d.). Critical Thinking, the Educated Mind, and the Creation of Critical Societies…Thoughts from the Past. *The Foundation For Critical Thinking*. https://www.criticalthinking.org/pages/critical-societies-thoughts-from-the-past/762.

Eldridge, G. (2010). *Enhancing critical thinking for all students*. The International Educator. https://www.tieonline.com/article/28/enhancing-critical-thinking-for-all-students.

Emir, S. (2009). Education faculty students' critical thinking disposition according to academic achievement. *Procedia Social and Behavioral* Sciences. *1*(1). pp. 2466-2469. https://doi.org/10.1016/j.sbspro.2009.01.433.

Ennis, R. H., Millman, J., & Tomko, T. N. (1985). *Cornell critical thinking tests level X & level Z: Manual*. Midwest Publications.

Ennis, R. H. & Weir, E. E. (1985). The Ennis-Weir critical thinking essay test: An instrument for teaching and testing. Pacific Grove, CA: Midwest Publications. https://www.academia.edu/1847582/The_Ennis-Weir_Critical_Thinking_Essay_Test_An_Instrument_for_Teaching_and_Testing.

Ennis, R. H. (1989). Critical Thinking and Subject Specificity: Clarification and Needed Research. *Educational Researcher. 18*(3). pp. 4–10. https://doi.org/10.3102/0013189X018003004.

Ennis, R. H. (2003). Critical thinking assessment. In D. Fasko (Ed.). *Critical thinking and Reasoning*. pp. 293–310. Cresskill, NJ: Hampton Press.

Erdmann, A. (2013). *How militaries learn and adapt: An interview with Major General H. R. McMaster*. McKinsey and Company http://www.mckinsey.com/insights/public_sector/how_militaries_learn_and_Adapt.

Esmaeilzad, S., Gavgani, V. Z., Zarei, A., & Familrouhany, S. A. A. (2022). Effect of evidence-based information management and practice training on librarians' critical thinking: A randomized educational trial. Journal of Librarianship and Information Science, 0(0). https://doi.org/10.1177/09610006221114648.

Facione, P. A. (1990). Critical Thinking: A Statement of Expert Consensus for Purposes of Educational Assessment and Instructions. *Research findings and recommendations*. California Academic Press.

Facione, P., Facione, N. & Giancarlo, C. (2000). The Disposition Toward Critical Thinking: Its Character, Measurement, and Relationship to Critical Thinking Skill. *Informal Logic. 20*. 10.22329/il.v20i1.2254. https://www.researchgate.net/publication/252896581_The_Disposition_Toward_Critical_Thinking_Its_Character_Measurement_and_Relationship_to_Critical_Thinking_Skill.

Facione, P. A. (2015). Critical Thinking: What It Is and Why It Counts. *Insight Assessment.* Measured Reasons LLC. http://www.insightassessment.com/Resources/Select-Tools-For-Teaching-For-and-About-Thinking/Critical-Thinking-What-It-Is-and-Why-It-Counts/Critical-Thinking-What-It-Is-and-Why-It-Counts.

Faravani, A. & Taleb, E. (2020). Teachers' use of Bloom's higher-order questions in class to augment EFL learners' listening comprehension and critical thinking ability. *Journal of Language and Cultural Education. 8*(2). pp. 94–113. https://doi.org/10.2478/jolace-2020-0015.

Farcis, F., Budi, G. & Wijayanti, E. (2022). Effect of Project-Based Learning and Science Literacy Ability on Critical Thinking Skills in Virtual Learning of the Thermodynamics Course. *JPPS (Jurnal Penelitian Pendidikan Sains). 12.* pp. 56-68. 10.26740/jpps. v12n1.p. 56-68. https://www.researchgate.net/publication/365777454_Effect_of_Project-Based_Learning_and_Science_Literacy_Ability_on_Critical_Thinking_Skills_in_Virtual_Learning_of_the_Thermodynamics_Course.

Fastiggi, W. (n.d.). Applying Bloom's Taxonomy to the Classroom. *Technology for Learners*. http://technologyforlearners.com/applying-blooms-taxonomy-to-the-classroom/.

Fisher, G., Hillison, J., & Haberichter, B. (2019). Competitive Advantage: More Than Just Technology. https://warroom.armywarcollege.edu/podcasts/competitive-advantage.

Fisher, R.A. (1925). *Statistical methods for research workers* (11th ed. rev.). Oliver and Boyd: Edinburgh.

Florence, C.D. (2014). A History of Critical Thinking as an Educational Goal in Graduate Theological Schools. *Christian Higher Education*, *13*:5. pp. 352-361, DOI: 10.1080/15363759.2014.949164. https://doi.org/10.1080/15363759.2014.949164

Forehand, M. (2005). Bloom's taxonomy: Original and revised. In M. Orey (Ed.), *Emerging perspectives on learning, teaching, and technology*. http://projects.coe.uga.edu/epltt.

The Foundation for Critical Thinking. (2019). *International Critical Thinking Test*. https://www.criticalthinking.org/pages/international-critical-thinkingtest/619.

Funk, P. E. (2020). Basic Officer Leader Training Policies and Administration. *Department of the Army, TRADOC Regulation 350-36*. https://adminpubs.tradoc.army.mil/regulations.html

Gelder, T.V. (2005) Teaching critical thinking: some lessons for cognitive science. *College Teaching*. *Winter*. *53*.1. https://www.reasoninglab.com/wp-content/uploads/2013/10/Tim-van-Gelder-Teaching-CT-Lessons-from-Cog-Sci.

Gerras, S. J. (2008). *Thinking Critically about Critical Thinking: A Fundamental Guide for Strategic Leaders*. U.S. Army War College. http://www.au.af.mil/au/awc/awcgate/army-usawc/crit_thkg_gerras.

Gerras, S. J. (2010). *Strategic Leadership Primer*. (3rd ed.). U.S. Army War College: Carlisle Barracks, PA. http://www.carlisle.army.mil/usawc/dclm/slp3

Gershon, M. (2015). *How to use Bloom's Taxonomy in the Classroom – The Complete Guide*. https://www.learningsciences.com/wp-content/uploads/2020/06/blooms_lookinside.

Giedrė Klimovienė, G., Urbonienė, J., and Barzdžiukienė, R. (2006). Developing Critical Thinking through Cooperative Learning. *Studies about languages*. https://www.kalbos.lt/zurnalai/09_numeris/11.

Glaser, E. M. (1972). *An Experiment in the Development of Critical Thinking. Teacher's College*. Columbia University. AMS Press.

Glass, G. (1976). Primary, Secondary, and Meta-analysis of Research. *Educational Researcher. 5*(10). pp. 3-8. https://doi.org/10.3102/0013189X005010003

Glass, G. V., McGaw, B., & Smith, M. L. (1981). *Meta-Analysis in Social Research*. Sage Publications.

Goodsett, M. (2020). Best practices for teaching and assessing critical thinking in information literacy online learning objects. *The Journal of Academic Librarianship*. *46*:5. https://www.sciencedirect.com/science/article/pii/S0099133320300665.

Guillot, M. W. (2004). Critical thinking for the military professional. *Air & Space Power Chronicles*. https://www.dau.mil/cop/pm/DAU%20Sponsored%20Documents/Critical%20Thinking%20For%20The%20Military%20Professional.htm.

Guolo, A., & Varin, C. (2015). Random-effects meta-analysis: The number of studies matters. *Statistical Methods in Medical Research. 26*(3). pp.1500 –1518.

https://doi.org/10.1177/0962280215583568.

Halawi, L.A., McCarthy, R.V. & Pires, S. (2009) An Evaluation of E-Learning on the Basis of Bloom's Taxonomy: An Exploratory Study. *Journal of Education for Business, 84*:6. pp. 374-380. DOI: 10.3200/JOEB.84.6.374-380

Halpern, D. F. (1998). Teaching critical thinking for transfer across domains: Dispositions, skills, structure training, and metacognitive monitoring. *American Psychologist. 53.* pp. 449–455. https://doi.org/10.1016/j.tsc.2009.02.001

Halpern, D. F. (2003). The “how” and “why” of critical thinking assessment. In D. Fasko (Ed.), *Critical thinking and reasoning: Current research, theory, and practice*. Hampton Press.

Halpern, D.F. (2010). *The Halpern critical thinking assessment*: Manual. Schuhfried.

Hansen, C., Steinmetz, H. & Block, J. (2022). How to conduct a meta-analysis in eight steps: a practical guide. *Manag Rev Q .72*. pp. 1–19. https://doi.org/10.1007/s11301-021-00247-4.

Hatcher, D. (2013). The Halpern Critical Thinking Assessment: A Review. *Inquiry: Critical Thinking Across the Disciplines. 28*. pp.18-23. 10.5840/inquiryct201328315. https://www.researchgate.net/publication/272705035_The_Halpern_Critical_Thinking_Assessment_A_Review.

Haynes, A., Lisic, E., Harris, K., Leming, K., Shanks, K., & Stein, B. (2015). “Using the Critical Thinking Assessment Test (CAT) as a Model for Designing Within-Course Assessments: Changing How Faculty Assess Student Learning.” Inquiry: Critical Thinking Across the Disciplines. 30(3). pp. 38–48.

doi:10.5840/inquiryct201530316.

Hayes, K.D. & Devitt, A.A. (2008), Classroom Discussions with Student-Led Feedback: *A Useful Activity to Enhance Development of Critical Thinking Skills. Journal of Food Science Education, 7*: 65–68. https://doi.org/10.1111/j.1541-4329.2008.00054.x.

Headquarters, Department of the Army. (2019). DA Pam 600-3. Commissioned Officer Development and Career Management. *Department of the Army Pamphlet 600-3*. https://armypubs.army.mil/ProductMaps/PubForm/Details.aspx?PUB_ID=1006180

Headquarters, Department of the Army. (2015). The Army Design Methodology. *Army Techniques Publication No. 5-0.1*. Washington, DC: Headquarters Department of the Army. https://armypubs.army.mil/epubs/DR_pubs/DR_a/pdf/web/atp5_0x1

Headquarters, United States Army Combined Arms Center. (2017). Operations (Field Manual 3-0). *Combined Arms Doctrine Directorate, United States Army Combined Arms Center*. Fort Leavenworth, KS: Headquarters Department of the Army. https://armypubs.army.mil/ProductMaps/PubForm/Details.aspx?PUB_ID=1003121

Headquarters, Department of the Army. (2011). Commander and Staff Officer Guide. *Army Tactics, Techniques, and Procedures No. 5-0.1*. Washington, DC: Headquarters Department of the Army. https://www.myarmypublications.com/general-army-subjects-c-43/attp-501-commander-and-staff-officer-guide

Hedges, L. V., Tipton, E., & Johnson, M. C. (2010). Erratum: Robust variance estimation in meta-regression with dependent effect size estimates. *Res. Synth. Methods 1*, 164–165.

Heft, I. E., & Scharff, L. F. V. (2017). Aligning Best Practices to Develop Targeted Critical Thinking Skills and Habits. *Journal of the Scholarship of Teaching and Learning.17*(3). July 2017. pp. 48-67. doi: 10.14434/josotl.v17i3.22600

Higgins, J. P. (2008). Commentary: Heterogeneity in meta-analysis should be expected and appropriately quantified. *International journal of epidemiology*, *37*(5), 1158-1160.

Higgins, J.P.T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M.J., Welch, V.A. (2021). Analyzing data and undertaking meta-analyses. *Cochrane Handbook for Systematic Reviews of Interventions version 6.2.* www.training.cochrane.org/handbook

Higgins, J.P.T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M.J., Welch, V.A. (editors). (2022). *Cochrane Handbook for Systematic Reviews of Interventions version 6.3 (updated February 2022).* Cochrane, 2022. https:// www.training.cochrane.org/handbook

Hines, R.L. (2018) Assembling the Pieces of a Systematic Review: A Guide for Librarians, edited by Margaret J. Foster and Sarah T. Jewell. *Medical Reference Services Quarterly, 37(*3). pp. 325-326. doi: 10.1080/02763869.2018.1477725.

Hitchcock, D. (2018). *Critical Thinking. The Stanford Encyclopedia of Philosophy*. Fall 2018 Ed., Edward N. Zalta (ed.).

https://plato.stanford.edu/archives/fall2018/entries/critical-thinking/

Hobaugh, C. F. (2010). Critical Thinking Skills: Do We Have Any? Critical Thinking Skills of Faculty Teaching Medical Subjects in a Military Environment. *U.S. Army Medical Department Journal*. pp. *48*–62. https://search-ebscohost-com.ezp.waldenulibrary.org/login.aspx?direct=true&db=tsh&AN=56544548&site=eds-live&scope=site

Holmes, N. G., Wieman, C. E., & Bonn, D. A. (2015). Teaching critical thinking. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(36), 11199–11204. https://doi.org/10.1073/pnas.1505329112

Huedo-Medina, T.; Sanchez-Meca, J.; Marin-Martinez, F.; and Botella, J. (2006). "Assessing heterogeneity in meta-analysis: Q statistic or I2 index?". *CHIP Documents*. https://opencommons.uconn.edu/chip_docs/19/?utm_source

Hunter, J. E., and Schmidt, F. L. (2000), Fixed Effects vs. Random Effects Meta-Analysis Models: Implications for Cumulative Research Knowledge. *International Journal of Selection and Assessment. 8*. pp. 275–292. doi:10.1111/1468-2389.00156. https://www.researchgate.net/publication/279899123_Fixed_Effects_vs_Random_Effects_Meta-Analysis_Models_Implications_for_Cumulative_Research_Knowledge

IBM SPSS Statistics for Windows. (2022). *IBM SPSS Statistics for Windows*. Version 28.01.0 (142) (IBM Corp., Armonk, N.Y., USA).

Innabi, H., & El Sheikh, O. (2007). The change in mathematics teachers' perceptions of critical thinking after 15 years of educational reform in Jordan. *Education*

Resources Information Center. (EJ748199). ERIC. https://doi.org/10.1007/s10649-005-9017-x

Insight Assessment. (2016). California Critical Thinking Skills Test User Manual and Resource Guide. *Insight Assessment*. San Jose, CA: The California Academic Press. https://www.insightassessment.com/article/california-critical-thinking-skills-test-cctst-2

Jackson, D., & Turner, R. (2017). *Power analysis for random-effects meta-analysis*. Res. Synth. Methods 8, 290–302. doi: 10.1002/jrs m.124. https://www.researchgate.net/publication/315802842_Power_analysis_for_random-effects_meta-analysis

Jackson, K., Kidder, K.L., Mann, S., Waggy II, W.H., Lander, N. & Zimmerman, S. R. (2000). *Raising the Flag: Implications of U.S. Military Approaches to General and Flag Officer Development*. Santa Monica, Calif.: RAND Corporation, RR-4347-OSD, 2020 https://www.rand.org/pubs/research_reports/RR4347.html.

Jelsma, O., Van Merrienboer, J. J. G., & Bijlstra, J. P. (1990). The ADAPT design model: Towards instructional control of transfer. *Instructional Science. 19. pp.* 89-120. https://www.researchgate.net/publication/225864400_The_ADAPT_design_model_Towards_instructional_control_of_transfer

Joint Operating Environment. (2016). *Joint Operating Environment 2035: The Joint Force in a Contested and Disordered World.* https://www.jcs.mil/Portals/36/Documents/Doctrine/concepts/joe_2035_july16

Johnson-Freese, J., & Kelley, K. P. (2017). Meaningful Metrics for Professional Military Education. *JFQ: Joint Force Quarterly*. *84*. pp. 65–71. https://ndupress.ndu.edu/Media/News/News-Article-View/Article/1038821/meaningful-metrics-for-professional-military-education/

Kagan, R. (2016). Emerging U.S. defense challenges and worldwide threats. *The Brookings Institution*. https://www.brookings.edu/articles/emerging-u-s-defense-challenges-and-worldwide-threats/

Kalimuddin, M. (2017). The Practical Application of Followership Theory in Mission Command. *Military Review, the Professional Journal of the U.S. Army Practical Online Exclusive*. Fort Leavenworth, KS: Army University Press. https://www.armyupress.army.mil/Journals/Military-Review/Online-Exclusive/2017-Online-Exclusive-Articles/Follower-ship-Theory/

Kassem, C. L. (2000, Winter). Implementation of a School-Wide Approach to Critical Thinking Instruction. *American Secondary Education*. *29*(2). pp. 26-36. http://www.jstor.org/stable/41064423.

Kaurin, P. S. (2017). Professional Military Education: What is it Good For? *The Strategy Bridge*. https://thestrategybridge.org/the-bridge/2017/6/22/professional-military-education-what-is-it-good-for.

Khadijeh Aghaei, K. & Mirzaei Rad, E. (2018). On the Interconnection between Bloom's Critical Thinking Taxonomy & Listening Comprehension Performance of Iranian English Foreign Language (EFL) Learners. *International Journal of English Language and Translation Studies*. *06*(03). pp. 22–31.

http://eltsjournal.org/archive/value6%20issue3/3-6-3-18.

Kim, T. W. & Mejia, S. (2019). From Artificial Intelligence to Artificial Wisdom: What Socrates Teaches Us. *Computer. Vol 52*:10. pp. 70-74. Doi: 10.1109/MC.2019.2929723. https://eds-b-ebscohost-com.ezp.waldenulibrary.org/eds/detail/detail?vid=1&sid=e1042cf5-688e-4ffd-8a47-ecd62776a6ef%40pdc-v-sessmgr05&bdata=JnNpdGU9ZWRzLWxpdmUm.

King, A. (1990). Reciprocal Peer-Questioning: A Strategy for Teaching Students How to Learn from Lectures. *The Clearing House, 64*(2). pp. 131-135. http://www.jstor.org/stable/30188588.

King, A. (1995). Designing the instructional process to enhance critical thinking across the curriculum: Inquiring minds really do want to know: Using questioning to teach critical thinking. *Teaching of Psychology*. *22* (1). pp. 13-17. https://web.s.ebscohost.com/ehost/pdfviewer/pdfviewer?vid=3&sid=d9dd72e5-28bf-4757-af88-459ea186413d%40redis

Knapp T. R. (2016). Why Is the One-Group Pretest-Posttest Design Still Used? *Clinical nursing research. 25*(5). pp. 467–472. https://doi.org/10.1177/1054773816666280

Kreitzberg, A. P. (2013). *Agile critical thinking: how to cope with change, complexity, and the unexpected*. Paper presented at PMI® Global Congress 2013—North America, New Orleans, LA. Newtown Square, PA: Project Management Institute

Ku, K.Y.L. (2009). Assessing students' critical thinking performance: Urging for measurements using the multi-response format. *Thinking Skills and Creativity. 4*(1). pp. 70-76. https://www.sciencedirect.com/science/article/abs/

pii/S1871187109000054.

Kuncel, N. R. (2011, January). *Measurement and meaning of critical thinking*. The report presented at the National Research Council's 21st Century Skills Workshop, Irvine, CA.

Laal, M. and Laal, M. (2012). Collaborative learning: What is it? *Procedia - Social and Behavioral Sciences. Vol. 31*.10.1016/j.sbspro.2011.12.092. https://www.sciencedirect.com/science/article/pii/S1877042811030217

Laerd Dissertation. (2012). *Reliability in research*. https://dissertation.laerd.com/reliability-in-research

Lage, M.J., Platt, G.J., & Treglia, M. (2000). Inverting the classroom: A gateway to creating an inclusive learning environment. J*ournal of Economic Education. 31*. pp. 30-43. 10.1080/00220480009596759. https://www.researchgate.net/publication/227450483_Inverting_the_Classroom_A_Gateway_to_Creating_an_Inclusive_Learning_Environment

Larkin, B. G., & Burton, K. J. (2008). Evaluating a case study using Bloom's taxonomy of education: The official voice of perioperative nursing. *AORN Journal. 88*(3). pp. 390-402. doi: http://dx.doi.org.ezp.waldenulibrary.org/10.1016/j.aorn.2008.04.020

Lasley II, T. J. (2016). Bloom's taxonomy. *Encyclopedia Britannica*. Encyclopedia Britannica, Inc. https://www.britannica.com/topic/Blooms-taxonomy

Lau, K. H., Lam, T. K., Booi, H. K., Nkhoma, M., & Richardson, J. (2018). Benchmarking higher education programs through alignment analysis based on the revised Bloom’s taxonomy. *Benchmarking. 25*(8). pp. 2828-2849. doi: http://dx.doi.org.ezp.waldenulibrary.org/10.1108/BIJ-10-2017-0286

Lederman, N. (2017). Re: What is the importance of Bloom's Taxonomy for the assessment and the strategies of the Cognitive Learning?. https://www.researchgate.net/post/What_is_the_importance_of_Blooms_Taxonomy_for_the_assessment_and_the_strategies_of_the_Cognitive_Learning/596e6b74f7b67ea7a34f6223.

Leana, C. R. (1985). A Partial Test of Janis' Groupthink Model: Effects of Group Cohesiveness and Leader Behavior on Defective Decision Making. *Journal of Management*, *11*(1), 5. https://doi.org/10.1177/014920638501100102

Lernier, J. (2010). Does the Digital Classroom Enfeeble the Mind? *New York Times*. September 16, 2010. http://www. nytimes.com/2010/09/19/magazine/19fob-essay-t.html?pagewanted=2

Lipsey, M. W., & Wilson, D. (2000). *Practical meta-analysis (applied social research methods)*. SAGE Publications.

Li, T., Higgins, J.P.T., Deeks, J.J. (editors) (2022). Chapter 5: Collecting data. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA (editors). *Cochrane Handbook for Systematic Reviews of Interventions* version 6.3 (updated February 2022). Cochrane, 2022. https://training.cochrane.org/handbook/current/chapter-05

Liu, O. L., Frankel, L., Crotts, K., & Roohr, K.C. (2014). Assessing Critical Thinking in Higher Education: Current State and Directions for Next-Generation Assessment.

Research Report ETS. RR–14-10. doi 10.1002/ets2.12009. https://onlinelibrary.wiley.com/doi/epdf/10.1002/ets2.12009

Liu, O.L., Frankel, L., & Roohr, K.C. (2014). Assessing Critical Thinking in Higher Education: Current State and Directions for Next-Generation Assessment. *ETS Research Reports*. doi: 10.1002/ets2.12009. https://files.eric.ed.gov/fulltext/EJ1109287.

Liu, O. L., Bridgeman, B., & Adler, R. M. (2012). Measuring Learning Outcomes in Higher Education: Motivation Matters. *Educational Researcher,* 41(9). pp. 352–362. https://doi.org/10.3102/0013189X12459679.

Liu, Y., & Xie, J. (2020). Cauchy combination test: A powerful test with analytic p-value calculation under arbitrary dependency structures. *J. Am. Stat. Assoc. 115.* pp. 393–402. doi:10.1080/01621459.2018.1554485.

López, F., Contreras, M., Nussbaum,M., Paredes, R., Gelerstein, D., Alvares, D.& Chiuminatto, P. (2023). Developing Critical Thinking in Technical and Vocational Education and Training. *Education Sciences*, *13*. p. 590. https://doi.org/10.3390/educsci13060590

López-Noguero, F. (2008). Towards a European area of lifelong learning: Evolution and development of Lifelong Learning in the European Union. *Social pedagogy: interuniversity journal. (15).* pp. 123-136. https://doi.org/10.1080/03050060802481496.

Lowther, A. & Mitchell, B. (2020). Professional Military Education needs more creativity, not more history. *War on the Rocks*.

https://warontherocks.com/2020/05/professional-military-education-needs-more-creativity-not-more-history/.

Lytell, M. C., Straus, S.G., Serena, C.C., Grimm, G.E., Doty III, J. L., Wenger, J.W., Abler, A.M., Naber, A. M., Grammich, C. A., & Fowler, E. S. (2017). *Assessing Competencies and Proficiency of Army Intelligence Analysts Across the Career Life Cycle*. RAND Corporation. RAND Corporation. RR-1851-A. https://www.rand.org/pubs/research_reports/RR1851.html

Macaskill, P., Walter, S. D., & Irwig, L. (2001). A comparison of methods to detect publication bias in meta-analysis. *Statistics in medicine, 20*(4). pp. 641–654. https://doi.org/10.1002/sim.698.

Machiavelli, N., Skinner, Q. & Price, R. (1988). *Machiavelli: The Prince*. Cambridge University Press. http://www.loc.gov/catdir/toc/cam031/88005048

Maggetti, M., Gilardi, F., & Radaelli, C. M. (2013). *Designing research in the social sciences*. SAGE Publications Ltd, https://doi.org/10.4135/9781473957664.

Magno, C. (2013). Assessing Students' Critical Thinking and Approaches to Learning. *The International Journal of Educational and Psychological Assessment*. *12.* https://www.researchgate.net/publication/277405273_Assessing_Students'_Critical_Thinking_and_Approaches_to_Learning

Maki, P. (2010). *Assessing for learning: Building a sustainable commitment across the institution* (2nd Ed.). Sterling, VA: Stylus. https://doi.org/10.4324/9781003443056

Marín-Díaz, V. (2015). Educational gamification. An alternative for creative teaching. *Digital Education Review.* https://www.researchgate.net/publication/290453359_Educative_gamification_An_alternative_to_creative_learning

Masadeh, M. (2012). Training, Education, Development: What is the Difference? European *Scientific Journal. 8* (10). https://www.researchgate.net/publication/279480522_Training_Education_Development_and_Learning_What_Is_The_Difference

Mayer, R. E. (2002). Meaningful learning versus rote learning. Theory into Practice. *The Ohio State University*. *41*: 4. Autumn. http://web.mit.edu/jrankin/www/teach_transfer/rote_v_meaning

Mayr, E. (2009) Darwin's Influence on Modern Thought. *Scientific American*. https://www.scientificamerican.com/article/darwins-influence-on-modern-thought1/.

Mayo-Wilson, E., Li, T.J., Fusco. N., Bertizzolo, L., Canner, J .K., Cowley, T., Doshi, P., Ehmsen, J., Gresham, G., Guo, N. J.A., Heyward, J., Hong, H., Pham, D., Payne, J.L., Rosman, L., Stuart, E.A., Suarez-Cuervo, C., Tolbert, E., Twose, C., Vedula, S., & Dickersin, K. (2017). Cherry-picking by trialists and meta-analysts can drive conclusions about intervention efficacy. *Journal of Clinical Epidemiology 2017a. 91.* pp. 95–110. https://www.sciencedirect.com/science/article/pii/S095435617307217?

Mayo-Wilson, E., Li, T., Fusco, N., & Dickersin, K. (2018). Practical guidance for using multiple data sources in systematic reviews and meta-analyses (with examples from the MUDS study). *Research Synthesis Methods. 9*: pp. 2–12. https://www.researchgate.net/publication/320573084_Practical_guidance_for_using_multiple_data_sources_in_systematic_reviews_and_meta-analyses_with_examples_from_the_MUDS_study_Using_multiple_data_sources_in_systematic_reviewsp.

McDade, S. A. (1995). Case study pedagogy to advance critical thinking. *Teaching Psychology, 22*(1). pp. 9-10. https://doi.org/10.1207/s15328023top2201_3

McLeod, S. A. (2017). Psychodynamic approach. *Simply psychology*. https://www.simplypsychology.org/psychodynamic.htmlp

McMaster, H.R. (2013). The Pipe Dream of Easy War. *The New York Times*. https://www.nytimes.com/2013/07/21/opinion/sunday/the-pipe-dream-of-easy-war.html

McPeck, J. E. (1981). *Critical thinking and education*. Oxford: Martin Robinson

Merrill, M. D. (2002). First Principles of Instruction. *Educational Technology Research and Development*, *50*(3). pp. 43–59. https://doi.org/10.1007/BF02505024.

Moffett, J., & Mill, A, C. (2014). Evaluation of the flipped classroom approach in a veterinary professional skills course. *Adv Med Educ Pract. 5*. pp.415–425. https://www.semanticscholar.org/reader/925a323bc19e5091ae63334b8b2cb92a98b3d921.

Mohd Salleh, K., & Sulaiman, N. L. (2019). The impact of organizational and professional development on human resource development practitioners in Malaysian organizations. *The Journal of Social Sciences Research. 5*(1), 683-689. https://www.researchgate.net/publication/331805274_The_Impact_of_Organizational_and_Professional_Development_on_Human_Resource_Development_Practitioners_in_Malaysian_Organizations/link/628b884339fa2170316826d7/download

Moore, T. J. (2011). Critical thinking and disciplinary thinking: A continuing debate. *Higher Education Research and Development*, 30(3). pp. 261–274. https://www.tandfonline.com/doi/full/10.1080/07294360.2010.501328?scroll=top&needAccess=true.

Monat, J.P & Gannon, T. F. (2017). Failures of Systems Thinking in U. S. Foreign Policy. *American Journal of Systems Science*. doi: 10.5923/j.ajss.20170501.01. http://article.sapub.org/10.5923.j.ajss.20170501.01

Morrison, S., & Free, K. W. (2001). Writing multiple-choice test items that promote and measure critical thinking. *Journal of Nursing Education*. 40. pp. 17-24. https://eds.s.ebscohost.com/eds/pdfviewer/pdfviewer?vid=2&sid=40405a10-b057-40ea-9060-d0cef83fae5c%40redis.

Morton, D. A., & Colbert-Getz, J. M. (2017). Measuring the impact of the flipped anatomy classroom: The importance of categorizing an assessment by Bloom's taxonomy. *Anatomical sciences education, 10*(2). pp. 170–175. https://doi.org/10.1002/ase.1635.

Mulcare, D. M., & Shwedel, A. (2017). Transforming Bloom's Taxonomy into Classroom Practice: A Practical Yet Comprehensive Approach to Promote Critical Reading and Student Participation. *Journal of Political Science Education. Vol. 13*(2). pp. 121–137. http://dx.doi.org/10.1080/15512169.2016.1211017

Muntersbjorn, M. (2003). Francis Bacon's Philosophy of Science: Machina intellectus and Forma indita. *Philosophy of Science. 70*. 10.1086/377395. 13(2). pp. 121–137. https://doi.org/10.1080/15512169.2016.121101

Munzenmaier, C., & Rubin, N. (2013). Bloom's taxonomy: what's old is new again. *The eLearning Guild*. https://www.elearningGuild.com.

Murawski, L. M. (2014) Critical Thinking in the Classroom…and Beyond. *Journal of Learning in Higher Education. 10*:1. pp. 25-30. Spr 2014. https://eric.ed.gov/?id=EJ1143316

Murphy, K. R. (2017). What inferences can and cannot be made on the basis of meta-analysis? *Human Resource Management Review. 27(1).* pp. 193-200. https://doi.org/10.1016/j.hrmr.2015.06.001

Murray, N. (2014). *The Role of Professional Military Education in Mission Command*. National Defense University Press. https://ndupress.ndu.edu/Media/News/Article/577475/the-role-of-professional-military-education-in-mission-command/.

Nolan, L. (1998). Descartes' Theory of Universals. *Philosophical Studies. 89*. pp. 161–180. https://doi.org/10.1023/A:1004228828029.

Nosich, G. M. (2009). *Learning to think things through: A guide to critical thinking across the curriculum*. Prentice-Hall.

OECD. (2013). Assessment of Higher Education Learning Outcomes. *Further Insights (Feasibility Study Report Vol. 3*). http://www.oecd.org/education/skills-beyond-school/AHELOFSReportVolume3.

Odierno, R. T. & McHugh, J. M. (2015). The Army Vision Strategic Advantage in a Complex World. Department of the Army. pp. 1-16. https://usacac.army.mil/publication/the_army_vision.

Onion, A. Sullivan, M., Mullen, M., & Zapata, C. (2018). Machiavelli. History. *A&E Television Networks*. https://www.history.com/topics/renaissance/machiavelli.

Page, M.J., Higgins, J.P.T., & Sterne, JA.C. (2021) Chapter 13: Assessing risk of bias due to missing results in a synthesis. In: Higgins, J.P.T., Thomas, J., Chandler, J., Cumpston, M., Li, T., Page, M.J., Welch, V.A. (editors). *Cochrane Handbook for Systematic Reviews of Interventions version 6.2* (updated February 2021). Cochrane, 2021. www.training.cochrane.org/handbook.

Patten, M.L. (2017). *Understanding Research Methods: An Overview of the Essentials* (10th ed.). Routledge. https://doi.org/10.4324/9781315213033.

Paul, R., Elder, L., & Bartell, T. (1997). *California Teacher Preparation for Instruction in Critical Thinking: Research Findings and Policy Recommendations*. California: Commission on Teacher Credentialing. https://eric.ed.gov/?q=California+Teacher+Preparation+for+Instruction+in+Critic al+Thinking%3a+Research+Findings+and+Policy+Recommendations&ft=on&id=ED437379.

Paul, R. (2005) The State of Critical Thinking Today. *The Foundation for Critical Thinking*. California: Foundations for Critical Thinking Press. doi: https://doi.org/10.1002/cc.193

Paul, R. & Elder, L. (2007a). White Paper Consequential Validity: Using assessment to drive instruction. *Foundation for Critical Thinking*. pp. 1-7. http://www.criticalthinking.org/pages/testing-and-assessment/594.

Paul, R. & Elder, L. (2007b). Critical Thinking Competency Standards: Standards, Principles, Performance Indicators and Outcomes with a critical thinking master rubric. *The Foundation for Critical Thinking*. pp. 1-11. https://www.criticalthinking.org/files/SAM_Comp%20Stand_07opt

Paul, R. & Elder, L. (2008). *The Miniature Guide to Critical Thinking Concepts and Tools*. California: Foundation for Critical Thinking Press. pp. 1-8. https://www.criticalthinking.org/files/Concepts_Tools.

Paul, R. & Elder, L. (2010). *The Miniature Guide to Critical Thinking Concepts and Tools*. Foundation for Critical Thinking Press. pp. 1-8. https://www.criticalthinking.org/files/Concepts_Tools.

Perkins, D.G. (2014a). The Future Army. *Center for Strategic & International Studies*. [Media]. Washington, DC.

Perkins, D.G. (2014b). The U.S. Army Operating Concept: Win in a complex world. *Training and Doctrine Command Pamphlet 525-3-1*. Fort Eustis, VA: U.S. Army

Training and Doctrine Command. https://usacac.army.mil/publication/us-army-operating-concept.

Perkins, D.G. (2016) *Developing Future Strategic Leaders*. [Lecture recording] Commandant's National Security Program. Army War College Distance Education Second Residence Course. Carlisle Barracks.

Petcharuk, R. & Namon Jeerungsuwan, N. (2021). Development of Information-assisted Problem-based Learning System in Smart Classroom Environment for Critical Thinking Skill Development. *Review of International Geographical Education*. *Vol. 11* No. 8. pp. 122-128. doi: 10.48047/rigeo.11.08.12. https://rigeo.org/menu-script/index.php/rigeo/article/view/1249.

Pierce, R. and Fox, J. (2012). Vodcasts and active-learning exercises in a "flipped classroom" model of a renal pharmacotherapy module. *Am J Pharm Educ 76*: 196. https://eds.s.ebscohost.com/eds/pdfviewer/pdfviewer?vid=2&sid=f78b7e2b-fd46-4c92-a879-e1a031dc59ea%40redis.

Pigott, T.D. (2019) Handling missing data. In Cooper, H., Hedges, L.V., Valentine, J.C. (eds). *The Handbook for Research Synthesis and Meta-analysis, 3rd ed*. pp. 399-416. New York Russell Sage. https://doi.org/10.7758/9781610448864.1.

Pleban, R. J., Tucker, J. S., Centric, J. H., Dlubac, M. D., & Richard L. Wampler, R. L. (2006). Assessment of the FY 05 Basic Officer Leader Course (BOLC) II Instructor Certification Program (ICP) and Single-Site Initial Implementation. *U.S. Army Research Institute for the Behavioral and Social Sciences*. Study

Report 2006-09. Alexandria, VA: Army Research Institute. https://apps.dtic.mil/dtic/tr/fulltext/u2/a460363

Pool-Funai, A. & Hansen, L. (2016). Weaving Critical Thinking with Experiential Learning in Public Administration. *American Society for Public Administration*. https://patimes.org/weaving-critical-thinking-experiential-learning-public-administration/.

Prawat, R. S. (2000). The Two Faces of Deweyan Pragmatism: Inductionism versus Social Constructivism. *Teachers College Record. 102*(4). pp. 805–840. https://doi.org/10.1111/0161-4681.00078.

Preville, Phillip. (n.d.). *The Professor's Guide to Using Bloom's Taxonomy*. https://tophat.com/teaching-resources/ebooks-and-guides/blooms-taxonomy-guide/.

Prober, C. G., and Khan, S. (2013). Medical education reimagined: A call to action. Acad Med 88. pp.1407–1410.

Public Law 114–95. (2015). *Every Student Succeeds Act of 2015*. Pub. L. No. 114–95. 129 Statute 1801–2192. https://congress.gov/114/plaws/publ95/PLAW-114publ95.

Qasrawi, R., & BeniAbdelrahman, A. (2020). The Higher and Lower-Order Thinking Skills (Hots and Lots) in Unlock English Textbooks (1St and 2Nd Editions) Based on Bloom's Taxonomy: An Analysis Study. *International Online Journal of Education & Teaching. 7*(3). pp. 744–758.

https://eds.s.ebscohost.com/eds/pdfviewer/pdfviewer?vid=2&sid=dc8f7abe-614b-43a3-9b01-4b0b02246e69%40redisp.

Quitadamo, I., & Kurtz, M. (2007). Learning to improve: Using writing to increase critical thinking performance in general education biology. *CBE Life Sciences Education, 6*(2). pp.140-154. Online. *11*(8). pp. 122–128. https://doi.org/10.48047/rigeo.11.08.12.

Radmehr, F. & Drake, M. (2019) Revised Bloom's taxonomy and major theories and frameworks that influence the teaching, learning, and assessment of mathematics: a comparison. *International Journal of Mathematical Education in Science and Technology*. *50*:6. pp. 895-920. doi: 10.1080/0020739X.2018.1549336. https://www.tandfonline.com/doi/epdf/10.1080/0020739X.2018.1549336?needAccess=true.

The Research Group on Socialism and Democracy. (2014). Critical Thinking and Class Analysis: Historical Materialism and Social Theory. *Socialism and Democracy Online*. http://sdonline.org/61/critical-thinking-and-class-analysis-historical-materialism-and-social-theory-2/.

Reiter, Bernd. (2012). Learning from Brazil and India: The difference that inclusion policies can make. *Government and International Affairs Faculty Publications*. *102*. http://scholarcommons.usf.edu/gia_facpub/102.

Renaud, R. & Murray, H. (2008). A comparison of a subject-specific and a general measure of critical thinking. *Thinking Skills and Creativity*. *3*. pp. 85-93. 10.1016/j.tsc.2008.03.005.

https://www.sciencedirect.com/science/article/abs/pii/S1871187108000096

Roberts, P., & Lawson, E. (2019). Contemporary Schools of War. The Future Conflict Operating Environment Out to 2030. *Royal United Services Institute for Defence and Security Studies Journal*. https://www.rusi.org/publication/occasional-papers/future-conflict-operating-environment-out-2030.

Robertson, J. F. & Rane-Szostak, D. (1996). Using dialogues to develop critical thinking skills: A practical approach. J*ournal of Adolescent & Adult Literacy*. *39*(7). pp. 552-556. https://www.proquest.com/docview/216923554/fulltextPDF/C13DD83EF96D4C95PQ/1?accountid=14872

Rodgers, C. (2002). Defining Reflection: Another Look at John Dewey and Reflective Thinking. *Teachers College Record. 104*(4). pp. 842–866. https://doi.org/10.1111/1467-9620.00181

Rodgers, R., Crosskerry, P., & Frank, B. (2015). The Challenges of Assessing Critical Thinking. *Higher Education Quality Council of Ontario*. [Webinar]. https://www.youtube.com/watch?time_continue=3215&v=9jYWclvOHfM&feature=emb_logo.

Roohr, K. C., & Burkander, K. (2020). Exploring Critical Thinking as an Outcome for Students Enrolled in Community Colleges. *Community College Review*, *48*(3), 330–351. https://doi.org/10.1177/0091552120923402.

Rosenbaum, B. (2003). The unconscious. *The Scandinavian Psychoanalytic Review. 26*. pp. 31-40. doi: 10.1080/01062301.2003.10592905. https://www.researchgate.net/publication/271820646_The_Unconscious.

Rosenblad, A. (2009). Introduction to Meta-Analysis by Michael Borenstein, Larry V. Hedges, Julian P.T. Higgins, Hannah R. Rothstein. *International Statistical Review*, *77*(3), 478–479. https://doi.org/10.1111/j.1751-5823.2009.00095_15.x.

Royalty, J. (1995). The generalizability of critical thinking: Paranormal beliefs versus statistical reasoning. J*ournal of Genetic Psychology*. *156*. 477–488. doi:10.1080/00221325.1995.9914838. https://eds.s.ebscohost.com/eds/pdfviewer/pdfviewer?vid=2&sid=e4dc7991-0217-47ef-9df9-dcd8a072830b%40redis

Ryan, R. (2016). Heterogeneity and subgroup analyses in Cochrane Consumers and Communication Group review: planning the analysis at protocol stage. *Cochrane Consumers and Communication Review Group*. http://cccrg.cochrane.org.

Sá, W. C., Stanovich, K. E., & West, R. F. (1999). The domain specificity and generality of belief bias: Searching for generalizable critical thinking skill. *Journal of Educational Psychology*. *91*. 497–510. doi:10.1037/0022-0663.91.3.497. https://eds.s.ebscohost.com/eds/pdfviewer/pdfviewer?vid=4&sid=3ed01771-95c6-40ad-8567-3bd81774668a%40redis

Sami, Joy & Umamakeswari, Arumugam. (2020). A Descriptive Analysis of Students' Learning Skills Using Bloom's Revised Taxonomy. *Journal of Computer Science, 16*, 183-193. 10.3844/jcssp.2020.183.193. https://www.researchgate.net/publication/340480210_A_Descriptive_Analysis_of_Students_Learning_Skills_Using_Bloom's_Revised_Taxonomy

Schafer, W. D. (1999). Methods, plainly speaking: An overview of meta-analysis. *Measurement and Evaluation in Counseling and Development, 32*(1), 43-61. https://eds.s.ebscohost.com/eds/detail/detail?vid=2&sid=8de71454-3245-4f1b-95a5-1107829b05ca%40redis&bdata= JkF1dGhUeXBlPXNoa WImc2l0ZT1lZHMtbGl2ZSZzY29wZT1zaXRl#AN=1999-13904-004&db=psyh

Schatz, S., Fautua, D.T., Stodd, J., & Reitz, E. A. (2017) The Changing Face of Military Learning. *Journal of Military Learning*. Fort Leavenworth, KS: The Army Press. http://www.armyupress.army.mil/Portals/7/journal-of-military-learning/Archives/jml-april-2017-Schatz-Fautua-Stodd-Reitz-Changing

Schmidt, F. L., & Hunter, J. E. (2014). *Methods of meta-analysis: Correcting error and bias in research findings*. Thousand Oaks, CA: Sage publications

Schumm, W.R; Webb, F. J.; Turek, D.E.; Jones, K. D. and Ballard, G. E. (2006). A comparison of methods for teaching critical thinking skills for U.S. Army Officers. *American Journal of Distance Education, 20*:1, 39-50, DOI: 10.1207/s15389286ajde2001_4. https://doi.org/10.1207/s15389286ajde2001_4

Scott, K. D. (2016). *The Joint Operating Environment 2035. The Joint Force in a contested and disordered world*. https://www.jcs.mil/Portals/36/Documents/Doctrine/concepts/joe_2035_july16.pdf?ver=2017-12-28-162059-917

Scriven, M. & Paul, R. (1987). Critical Thinking Defined. *8th Annual International Conference on Critical Thinking and Education Reform*, Summer. https://pct.libguides.com/c.php?g=134053&p=877568

Secolsky, C., & Denison, D. B. (Eds.). (2011). *Handbook on measurement, assessment, and evaluation in higher education*. https://search-proquest-com.ezp.waldenulibrary.org/legacydocview/EBC/957428?accountid=14872.com

Senge, P. M. (2006). *The Fifth Discipline: The Art and Practice of The Learning Organization*. New York: Doubleday

Serrat O. (2017) Critical Thinking. *In: Knowledge Solutions*. Springer, Singapore. https://doi.org/10.1007/9 78-981-10-0983-9_125

Sideeg, A. (2016). Bloom's Taxonomy, Backward Design, and Vygotsky's Zone of Proximal Development in Crafting Learning Outcomes. *International Journal of Linguistics*. *Vol. 8,* 158-186. 10.5296/ijl.v8i2.9252. https://www.researchgate.net/figure/Vygotskys-Zone-of-Proximal-Development_fig3_301684299

Siegel, H. (1980). Critical Thinking as an Educational Ideal. *The Educational Forum, 45*(1), 7-23. doi: http://dx.doi.org/10.1080/00131728009336046.

Simon, M.K. (2011). Assumptions, limitations, and delimitations. *Dissertation and scholarly research: Recipes for success.* Seattle, WA: Dissertation Success, LLC

Şimşek, Y. (2017). The Evaluation of the Application of Transported Education by Teachers. *International Journal of Educational Research Review*. 2(1), 41– 48. DOI: https://doi.org/10.24331/ijere.309972.

Shana, Z. (2009). Learning with technology: Using discussion forums to augment a traditional-style class. *Educational Technology & Society*, *12* (3), 214–228.

https://eds.s.ebscohost.com/eds/pdfviewer/pdfviewer?vid=2&sid=ac2b3ac7-4b17-4601-9927-d324c6fc6ec1%40redis

Shelby, L. & Vaske, J. (2008). Understanding Meta-Analysis: A Review of the Methodological Literature. *Leisure Sciences - LEISURE SCI*, *30*, 96-110. 10.1080/01490400701881366. https://eds.s.ebscohost.com/eds/pdfviewer/pdfviewer?vid=2&sid=6c60c69f-a87a-4ddc-8c17-ab88b901d415%40redis

Shermis, M. D., & Burstein, J. C. (2003). Automated essay scoring: A cross-disciplinary perspective. Hillsdale, NJ: Erlbaum: Routledge

Shively, K., Stith, K. M., & Rubenstein, L. D. (2018). Measuring what matters: Assessing creativity, critical thinking, and the design process. *Gifted child today. 41*(3). 149–158. https://doi.org/10.1177/1076217518768361

Smith, G. (2002). Are there domain-specific thinking skills? *Journal of Philosophy of Education. 36*, 207–227. doi:10.1111/1467-9752.00270. https://eds.s.ebscohost.com/eds/pdfviewer/pdfviewer?vid=2&sid=dd5aad63-34a6-43a9-8979-0eb1e03ec687%40redis

Smith, T. (2019). Army's long-awaited Iraq war study finds Iran was the only winner in a conflict that holds many lessons for future wars. *Army Times*. https://www.armytimes.com/news/your-army/2019/01/18/armys-long-awaited-iraq-war-study-finds-iran-was-the-only-winner-in-a-conflict-that-holds-many-lessons-for-future-wars/

Society for Human Resource Management (SHRM). (2019). *The Global Skills Shortage Bridging the Talent Gap with Education, Training, and Sourcing*.

https://www.shrm.org/hr-today/trends-and-forecasting/research-and-surveys/documents/shrm%20skills%20gap%202019

Soozandehfar, S.M., & Adeli, M.R. (2016). A Critical Appraisal of Bloom's Taxonomy. *American Research Journal of English and Literature. 2*. DOI:10.21694/2378-9026.16014. https://www.arjonline.org/papers/arjel/v2-i1/14

Southworth, J. (2022). Bridging critical thinking and transformative learning: The role of perspective-taking. *Theory and Research in Education, 20*(1), 44–63. https://doi.org/10.1177/14778785221090853

Spicer, K. L. & Hanks, W. E. (1995). *Multiple measures of critical thinking skills and predisposition in assessment of critical thinking*. Paper presented at the annual meeting of the Speech Communication Association: San Antonio, TX. (Eric Document Reproduction Services No. ED 391 185). https://eric.ed.gov/?id=ED391185

Straus, S. G., Galegher, J., Shanley, M. G., & Moini, J. S. (2006). *Improving the Effectiveness of Distributed Learning: A Research and Policy Agenda*. www.rand.org. http://www.rand.org/pubs/occasional_papers/OP156.html

Straus, S. G., Shanley, M. G., Crowley, J. C., Yeung, D., Bana, S. H., & Leuschner, K.J. (2013). Developing Army Leaders Lessons for Teaching Critical Thinking in Distributed, Resident, and Mixed-Delivery Venues. *RAND Arroyo Center*. Santa Monica. CA: Rand Publications. https://www.rand.org/pubs/research_reports/RR321.html

Straus, S. G., Shanley, M. G., Crowley, J.C., Yeung, D., S. H. & Leuschner, K. J. (2014). *Developing Army Leaders: Lessons for Teaching Critical Thinking in Distributed, Resident, and Mixed-Delivery Venues*. Santa Monica, CA: RAND Corporation. https://www.rand.org/pubs/research_reports/RR321.html

Straus, S. G., Krueger, T. C., Geoffrey E. Grimm, G.E., & Giglio, K. (2018). *Malleability and Measurement of Army Leader Attributes: Personnel Development in the U.S.* Army. Santa Monica, CA: RAND Corporation, RR-1583-A. https://www.rand.org/pubs/research_reports/RR1583

Sternberg R.J. (1990). *Thinking styles: Keys to understanding performance*. Phi Delta Kappan. 366-371

Stone, A. J. (2017). Critical Thinking Skills in USAF Developmental Education. Air & *Space Power Journal*. *31*(2). pp. 1-16. https://www.airuniversity.af.edu/Portals/10/ASPJ/journals/Volume-31_Issue-2/F-Stone.

Strohm, S. M., & Baukus, R. A. (1995). Strategies for fostering critical thinking skills. *Journalism and Mass Communication Educator, 50* (1), 55-62. https://eds.s.ebscohost.com/eds/pdfviewer/pdfviewer?vid=2&sid=ee6dd009-14e0-420e-a263-0e26ca23e064%40redis.

Summary of the Irregular Warfare Annex to the National Defense Strategy 2020. (2020). *U.S. Department of Defense*. https://media.defense.gov/2020/Oct/02/2002510472/-1/-1/0/Irregular-Warfare-Annex-to-the-National-Defense-Strategy-Summary

Sumner, W. G. (1906). *Folkways - A Study of the Sociological Importance of Usages, Manners, Customs, Mores, and Morals*. http://www.gutenberg.org/files/24253/24253-h/24253-h.htm

Sutton, H. (2019), Integrate different generational needs for a strong adult learner experience. *Enrollment Management Report*, *23*: 6-7. https://doi.org/10.1002/emt.30523.

Taherdoost, H. (2016). Sampling Methods in Research Methodology; How to Choose a Sampling Technique for Research. *International Journal of Academic Research in Management. 5*. 18-27. 10.2139/ssrn.3205035. https://www.researchgate.net/publication/319998246_Sampling_Methods_in_Research_Methodology_How_to_Choose_a_Sampling_Technique_for_Research

Terenzini, P. T., Springer, L., Pascarella, E. T., & Nora, A. (1995). Influences affecting the development of students' critical thinking skills. *Research in Higher Education. 36*(1). 23-39. https://files.eric.ed.gov/fulltext/ED372666

Theokas, C. (2010). *Shut out of the Military: Today's High School Education Doesn't Mean You're Ready for Today's Army*. Washington, DC: Education Trust http://www.edtrust.org/sites/edtrust.org/files/publications/files/ASVAB_4

Thomas, L. (2022). *Cross-Sectional Study | Definition, Uses & Examples.* Scribbr. https://www.scribbr.com/methodology/cross-sectional-study/

Thompson, A.R., & O'Loughlin, V.D. (2015). The Blooming Anatomy Tool (BAT): A discipline specific rubric for utilizing Bloom's taxonomy in the design and evaluation of assessments in the anatomical sciences. *Anat Sci Educ. 8*:493–501. https://doi.org/10.1002/ase.1507

Thurmond, V. A. (2001). The Holism in Critical Thinking: A Concept Analysis. *Journal of Holistic Nursing. 19*(4). 375–389. https://doi.org/10.1177/089801010101900406

Tierney, N.J. & Cook, D.H. (2018) Expanding tidy data principles to facilitate missing data exploration, visualization, and assessment of imputations. *Computation (stat.CO).* https://doi.org/10.48550/arXiv.1809.02264

Tremblay, K., Lalancette, D., & Roseveare, D. (2012). Assessment of Higher Education Learning Outcomes. *Design and Implementation (Feasibility Study Report Vol. 1).* http:// www.oecd.org/education/skills-beyond-

Turner, R. M., Bird, S. M., & Higgins, J. P. (2013). The impact of study size on meta-analyses: examination of underpowered studies in Cochrane reviews. *PloS one*, *8*(3), e59202. https://doi.org/10.1371/journal.pone.0059202.

Underwood, M. K., & Wald, R. L. (1995). Conference-style learning: A method for fostering critical thinking with heart. *Teaching of Psychology*. *22*:1, 17-21, DOI: 10.1207/s15328023top2201_6

The University of Chicago Chronicle. (1999). Bloom, influential education researcher. *The University of Chicago Chronicle. 19*:1. https://chronicle.uchicago.edu/990923/bloom.shtml

Uma, D., Thenmozhi, S., & Hansda, R. (2017). Analysis on Cognitive Thinking of an Assessment System Using Revised Bloom’s Taxonomy. *2017 5th IEEE*

International Conference on MOOCs, Innovation, and Technology in Education (MITE), MOOCs, Innovation and Technology in Education (MITE), 2017 5th IEEE International Conference on, MITE, 152–159. https://doi-org.ezp.waldenulibrary.org/10.1109/MITE.2017.00033

Umscheid, C. A. (2013). *A Primer on Performing Systematic Reviews and Meta-analyses, Clinical Infectious Diseases*. *57*(5). https://doi.org/10.1093/cid/cit333

Underwood, M. K., & Wald, R. L. (1995). Conference-style learning: A method for fostering critical thinking with heart. *Teaching Psychology, 22*(1), 17-21. https://eds.s.ebscohost.com/eds/pdfviewer/pdfviewer?vid=2&sid=c364e23e-febe-4083-bd07-7324d91c1bfa%40redis

United States Army Training and Doctrine Command (TRADOC) (2018). *TRADOC Pamphlet 350-70-7 Army Learning, Army Education Processes*. Department of the Army Headquarters, United States Army Training and Doctrine Command. https://adminpubs.tradoc.army.mil/regulations/TR350-70

United States Army Training and Doctrine Command. (2018). The U.S. Army in Multi-Domain Operations 2028. *Department of the Army Headquarters, TRADOC Pamphlet 525-3-1*. https://adminpubs.tradoc.army.mil/pamphlets/TP525-3-1

The University of Chicago Chronicle. (1999). Bloom, influential education researcher. *The University of Chicago Chronicle. 19:1*. https://chronicle.uchicago.edu/990923/bloom.shtml

Unrau, J. N. (1997). *Thoughtful teachers, thoughtful learners: A guide to helping adolescents think critically*. Ontario: Pippin Publishing Corporation

Valentine, J. C., Pigott, T. D., & Rothstein, H. R. (2010). How Many Studies Do You Need?: A Primer on Statistical Power for Meta-Analysis. *Journal of Educational and Behavioral Statistics, 35*(2), 215–247. https://doi.org/10.3102/1076998609346961.

Veldkamp, A.; van de Grint, L.; Knippels, M. P.J.; & van Joolingen, W.R. (2020). Escape education: A systematic review on escape rooms in education. *Educational Research Review. 31*. https://doi.org/10.1016/j.edurev.2020.100364.

Vogler, K. E. (2002). The impact of high-stakes, state-mandated student performance assessment on teachers' instructional practices. *Education. 123*(1). 39–55. https://eds.s.ebscohost.com/eds/pdfviewer/pdfviewer?vid=2&sid=ade29a78-9955-4ddb-8c08-95ddd81a0054%40redis.

Wade, C. & Tavris, C. (1987). *Using Writing to Develop and Assess Critical Thinking: Teaching of Psychology*. Psychology (1st ed.). 22(1): 24-28. New York: Harper. IN. https://doi.org/10.1207/s15328023top2201_8.

Walker Center for Teaching and Learning. (n.d.) Critical Thinking and Problem-solving. *The University of Tennessee at Chattanooga*. Online Resources for Faculty. https://www.utc.edu/walker-center-teaching-learning/teaching-resources/ct-ps.php

Walker, E., Hernandez, A. V., & Kattan, M. W. (2008). Meta-analysis: Its strengths and limitations. *Cleveland Clinic Journal of Medicine*, 75(6), 431–439. doi: 10.3949/ccjm.75.6.431. https://pubmed.ncbi.nlm.nih.gov/18595551/.

Wallace, E. D., & Jefferson, R. N. (2015). Developing Critical Thinking Skills: Assessing the Effectiveness of Workbook Exercises. *Journal of College Teaching &*

Learning. *12*(2), 101–108. https://eric.ed.gov/?id=EJ1061312

Wang, X., & Cheng, Z. (2020). Cross-Sectional Studies: Strengths, Weaknesses, and Recommendations. *Chest*, *158*(1S), S65–S71. https://doi.org/10.1016/j.chest.2020.03.012.

Watson, G. (1980). *Watson-Glaser critical thinking appraisal*. San Antonio, TX: Psychological Corporation. https://www.pearson.com/content/dam/one-dot-com/one-dot-com/global/Files/efficacy-and-research/reports/Watson-Glaser_One_Page_Summary.

Weeks, S. (2019). Francis Bacon's doctrine of idols: A diagnosis of 'universal madness.' *The British Journal for the History of Science, 52*(1), 1-39. doi:10.1017/S0007087418000961. https://www.proquest.com/docview/2188042580?accountid=14872

Westen, D., & Rosenthal, R. (2003). Quantifying construct validity: Two simple measures. *Journal of Personality and Social Psychology 84*. No. 3: 608-618. https://doi.org/10.1037/0022-3514.84.3.608.

Williams, T. M. (2013). Education for Critical Thinking. *Military Review*. https://apps.dtic.mil/docs/citations/ADA576054.

Wilson, D. B. (2010). Practical Meta-Analysis. *Evaluators' Institute.* July 16-17. http://slidegur.com/doc/5512009/presentation.

Wineburg, S., & Schneider, J. (2009). Was Bloom's Taxonomy Pointed in the Wrong Direction? *The Phi Delta Kappan, 91*(4), 56-61. http://www.jstor.org/stable/25594682.

Winkler, J.D., Marler, T., Marek N. Posard, M. N., Cohen, R. S., and Smith, M.L (2019) *Reflections on the Future of Warfare and Implications for Personnel Policies of the U.S. Department of Defense*. Santa Monica, Calif.: RAND Corporation. PE-324-OSD. As of January 12, 2021: https://www.rand.org/pubs/perspectives/PE324.html.

White, N. (2014). Organizing for War: Overcoming Barriers to Whole-of-Government Strategy in the ISIL Campaign. *The Small Wars Journal*. 12-28. http://smallwarsjournal.com/jrnl/art/organizing-for-war-overcomingbarriers-to-whole-of-government-strategy-in-the-isil-campaign.

Wong, L. Dr. & Gerras, S. J. Dr. (2013). "*Changing Minds In The Army: Why It Is So Difficult and What To Do About It."* Monographs, Books, and Publications. 515. https://press.armywarcollege.edu/monographs/515.

Wood, J. (2008). Methodology for dealing with duplicate study effects in a meta-analysis. *Organ. Res. Methods 11*, 79–95. doi:10.1177/1094428106296638. https://journals.sagepub.com/doi/epdf/10.1177/1094428106296638.

Wood W. & Eagly, A. H. (2009). Advantages of certainty and uncertainty. In Cooper H., Hedges L. V., Valentine J. C. (Eds.). *The Handbook of Research Synthesis and Meta-analysis* (2nd ed., pp. 455–472). New York, NY: Russell Sage Foundation.

Wright, T. (2021). *Global security challenges and strategy*. Brookings Institute. https://www.brookings.edu/testimonies/global-security-challenges-and-strategy/

Yasmin, K. (2020) How much difference can we make? Assessing the change in students' critical thinking in a private political science program in Egypt. *Journal* *of College Reading and Learning. 50*:4. 244-266. DOI: 10.1080/10790195.2020.1712271.

Yousef, W. (2021). *An assessment of critical thinking in the Middle East: Evaluating the effectiveness of special courses interventions*. PLoS ONE 16(12): e0262088. https://doi.org/10.1371/journal.pone.0262088.

Zaidi, N.B., Hwang, C., Scott, S., Stallard, S., Purkiss, J., & Hortsch, M. (2017). Climbing Bloom's Taxonomy pyramid: Lessons from a graduate histology course. *Anat Sci Educ. 2017, 10*:456–464. https://deepblue.lib.umich.edu/bitstream/handle/2027.42/138235/ase1685_am.;jsessionid=AA82CBCD5ADCD9083059FDE49563FE58?sequence=1.

Zapalska, A.M., McCarty, M.D., Young-McLear, K., & White, J. (2018). Design of assignments using the 21st-century Bloom's revised taxonomy model for development of critical thinking skills. *Problems and Perspectives in Management, 16*(2), 291-305. doi:10.21511/ppm.16(2).2018.27. https://doaj.org/article/fbdf776f6fb5454eaef9e2e995389171.

Zayapragassarazan, Z & Chacko, T.V. (2019). A gap analysis of critical thinking skills and attitude toward critical thinking among interns. *Int J Health Allied Science. 8*:193-6. https://eric.ed.gov/?id=ED598368.

Zhu, Q., and Carriere, K. (2018). Detecting and correcting for publication bias in meta-analysis – A truncated normal distribution approach. *Statistical Methods in Medical Research*. *2018; 27*(9):2722-2741. doi:10.1177/0962280216684671. https://journals.sagepub.com/doi/epub/10.1177/0962280216684671.

Zientek, L.; Yetkiner, Z.; Özel, S.; & Allen, J. (2012). Reporting Confidence Intervals and Effect Sizes: Collecting the Evidence. *Career and Technical Education Research Journal*. *37*. 277-295. 10.5328/cter37.3.277. https://eds.s.ebscohost.com/eds/pdfviewer/pdfviewer?vid=2&sid=7fff7f81-1625-430a-98c4-bb2876c6bc60%40redis.