Spring 2-28-2023

Analyzing the Cognitive Complexity of the Questions Contained on Assessments of College and Career Readiness for Grades 6-12

Sean D. Cronin
Seton Hall University, sean.cronin@student.shu.edu

Analyzing the Cognitive Complexity of the Questions Contained on Assessments of College and Career Readiness for Grades 6-12

Sean D. Cronin

Dissertation Committee

Christopher H. Tienken, EdD, Mentor

Mark Jarmon, EdD

David B. Reid, PhD

Submitted in partial fulfillment of the

requirements for the degree of

Doctor of Education

In the Department of Education Leadership, Management and Policy

Seton Hall University

2023

SETON HALL UNIVERSITY
1856

COLLEGE OF EDUCATION AND HUMAN SERVICES
DEPARTMENT OF EDUCATION LEADERSHIP MANAGEMENT & POLICY

APPROVAL FOR SUCCESSFUL DEFENSE

Sean D. Cronin has successfully defended and made the required modifications to the text of the doctoral dissertation for the Ed. D.during this Spring Semester, 2023.

DISSERTATION COMMITTEE

Dr. Christopher Tienken	2/28/2023
Mentor	Date
Dr. David Reid	2/28/2023
Committee Member	Date
Dr. Mark Jarmon	2/28/2023
Committee Member	Date

Abstract

This convergent, parallel, mixed-methods study with qualitative and quantitative content analysis methods was conducted to identify what type of thinking is required by the College and Career Readiness Assessment (CCRA+) by (a) determining the frequency and percentage of questions categorized as higher-level thinking within each cell of Hess' Cognitive Rigor Matrix and (b) determining in what ways are the skills identified as essential for success in the workplace by global organizations assessed on the CCRA+? The qualitative method consisted of a content analysis of the language of the analyzed assessment prompts, followed by deductive coding, and culminated with categorizing the depth and type of thinking required based on Hess' Cognitive Rigor Matrix, a framework that superimposes Bloom's Revised Taxonomy and Webb's Depth of Knowledge. To ensure reliability in the coding of the assessment prompts, the double-rater read behind the consensus model was employed. The results of the study found 70 percent of the analyzed questions required level 2 thinking according to Webb's Depth of Knowledge and understanding per Bloom's Revised Taxonomy. The study found over 82% of the selected-response questions required lower-level thinking from both Webb's Depth of Knowledge and Bloom's Taxonomy. All (100%) of the performance task assessment prompts that were analyzed were found to require Level 4 thinking on Webb's Depth of Knowledge and higher-level thinking according to Bloom's Revised Taxonomy. The results suggest the format of instruction and, ultimately, the format of assessment prompts are essential to develop and assess the development of student's critical thinking.

Keywords/terms: Bloom's Revised Taxonomy, Cognitive Complexity, Functional Fixedness, Global Competitiveness, Hess' Cognitive Rigor Matrix, Higher-order thinking, Reflective

Thinking, Selected-Response Questions, Standardized Assessment, Webb's Depth of Knowledge

DEDICATION AND ACKNOWLEDGEMENTS

I am beyond fortunate to be filled with love and support throughout the entirety of my life. As far as I can remember, my parents have distilled an unwavering belief in me that has allowed me to pursue my personal and professional goals, knowing I can always count on their encouragement and wisdom at all times. The process of this doctoral journey is a continuation of the support I have received from my brother, Brian, who questioned not if, but when I would earn my Doctorate of Education. I am incredibly thankful to know that you all are always there for me, no matter the need.

Many describe the journey of earning a doctoral degree to be filled with perseverance and resilience. The person who truly displayed these qualities over my time enrolled at Seton Hall University and always is my wife, Lea. I would describe taking classes and researching a topic I am interested in a hobby. The obligations and extra work that fell on you are something I am more than understanding and appreciative of. Thank you for your selflessness and dedication to raising our children, showering them with love and affection on days when I was in class and nights I was writing or completing coursework. No matter how exhausting the day was, you always took time ask about the program and to support me through this time. This support and your perspective is something I cherish. To my children, Cruz and Kaia, I hope what I gleaned from the research and writing of this program will benefit you both in your learning experiences both in and outside of school, with the dream of you keeping your curiosity and joyful spirits intact. The twinkle in your eyes and your sense of wonder is always an inspiration.

To my mentor, Dr. Tienken, one of the driving factors that compelled me to enroll in a doctoral program was reading *The School Reform Landscape: Fear, Myths, and Lies* when I was enrolled in graduate school. Your knowledge of the history of education, combined with your wisdom, is a gift to everyone who has the opportunity to learn from you. The generosity of your time, encouragement, and support were all greatly needed throughout this process. Our conversations always were thought-provoking and led me to look deeper into the process of learning, curricula, and the education system. Thank you for strengthening my thinking, more specifically, my process of critiquing educational practices and systems.

Dr. Reid, I am grateful for the opportunity to continue to learn from you beyond your classes in the doctoral program at Seton Hall University. Your efficient feedback, as well as the ability to determine only necessary information combined with your patience, has been a skill set I have admired since first having you as a professor. Dr. Jarmon, thank you for your prompt feedback, with an eye for details and clarity. Your expertise in education has been of tremendous assistance when fine tuning my writing. Dr. Creighton, thank you for always providing encouragement along with careful feedback to move my dissertation along. Dr. Solis-Stovall and Dr. Sydoruk, thank you both for taking the time to meet during the holiday season to discuss the methodology of this study, provide feedback, and code questions on your own time. I am more than thankful for the time all of you have spent to be a part of and improve my dissertation. Being able to count on you all to spend time improving my work in addition to all of your other obligations is something I am indebted to.

Table of Contents

List of Tables

List of Figures

Chapter I

Introduction

Background

Preparing students to be contributing members of global democratic societies and successful in the knowledge economy are two responsibilities of K-12 schools. Students entering kindergarten in the United States in the year 2022 are expected to begin to retire around the year 2082. Although the specific jobs those kindergarten students will attain cannot yet be determined, the skillset to be successful in the future world of work continues to receive attention in the mainstream press and private economic organizations. International organizations, such as the World Economic Forum (WEF), International Business Machines (IBM), and International Institute for Management Development (IMD) provide rankings and feedback from leading corporations and Chief Executive Officers (CEOs) on the skills determined essential to perform and be competitive in the job market.

Recent reports have identified innovation as a key characteristic among economically high-performing corporations. The 2021 CEO Study reported by IBM discussed “open innovation” as a “key theme that emerged” (p. 6). Innovation was also cited as an important factor. The 2016 Payscale and Future Workplace, Workforce-Skills Preparedness Report, which surveyed approximately 64,000 managers, noted the following skills are sought after by employers: critical thinking, problem solving, attention to detail, communication, writing proficiency, public speaking, and interpersonal skills/teamwork.

The World Competitiveness Rankings for the year 2021, published by IMD, listed innovation as the first of four key characteristics influencing economic success. The IMD (2021) claimed that reviving and transforming the innovation ecosystem will be important to economic strength and success. Additionally, the WEF (2020) suggested skills and dispositions such as "complex problem-solving," "critical thinking and analysis," and "creativity, originality and initiative" as three of the top ten skills required for the jobs of 2025. They listed the most important skill as "Analytical thinking and innovation" (WEF, 2020, p. 36). Organizations such as the WEF and IMD cite education as an important factor in the development of the skills they state will be necessary to be economically successful in the future.

As we embark on the Fourth Industrial Revolution, schools play a critical role in developing citizens who are prepared with the skills identified in the previous paragraphs to contribute and ultimately be economically viable. The *Fourth Industrial Revolution* is a term developed by Klaus Schwab, the Founder and Executive Chairman of the WEF. Defined as "a fusion of technologies that is blurring the lines between the physical, digital and biological spheres" (Schwab, 2016, para. 2). The Fourth Industrial Revolution provides "the potential to raise global income levels and improve the quality of life for populations around the world" (Schwab, 2016, para. 6). As technologies expand, schools must prepare all students for the skillsets needed to be successful and employable. McGinnis (2020) explains the Fourth Industrial Revolution as "a fusion of advances in artificial intelligence, robotics, the Internet of Things, 3D printing, genetic engineering, quantum computing, and other technologies (para. 1)". When describing the speed of implementation of the Fourth Industrial Revolution, Schwab (2016) wrote:

> When compared with previous industrial revolutions, the Fourth is evolving at an exponential rather than a linear pace. Moreover, it is disrupting almost every industry in every country. And the breadth and depth of these changes herald the transformation of entire systems of production, management, and governance. (para. 3)

With such changes to the knowledge economy, the way student skills are assessed is imperative to match what is needed in our current industrial revolution.

Standardized Testing

Standardized testing has played a role in the educational system beginning in the 19th Century. Tienken (2020) noted the origin of standardized assessments writing, "Large-scale use of commercial standardized tests in the United States began around 1918 with the introduction of the U.S. Army's Alpha and Beta tests of intelligence" (p. 73). The Elementary and Secondary Education Act (ESEA) of 1965 was the first federal legislation to require standardized testing by every Title 1 public school in the United States. The No Child Left Behind Act (NCLB) of 2001 broadened the mandate of testing to every public school in the United States. A provision in NCLB claimed that standardized tests were necessary for "improving and strengthening accountability, teaching, and learning by using State assessment systems designed to ensure that students are meeting challenging State academic achievement and content standards and increasing achievement overall" (No Child Left Behind Act of 2001, 2002, p. 16).

The Every Student Succeeds Act (ESSA) replaced NCLB in 2015. ESSA continued to require states to administer standardized tests in order to receive federal education funding. To be in compliance with the assessment component of ESSA every state must test students in reading and math once a year in Grades 3 through 8, and once in high school. Students must also be tested in science between grades K-5, 6-8, and 9-12.

The New Jersey Department of Education (NJDOE) required three standardized assessments for the 2021-2022 school year: The Start Strong Assessment; New Jersey Student Learning Assessment (NJSLA) to satisfy the ESSA testing requirements; and New Jersey high school exit exam, the Graduation Proficiency Assessment (NJGPA), in Grade 11. New Jersey is one of only 11 states that still requires a high school exit exam.

Every public school in the state of New Jersey is required to administer the NJSLA to satisfy the ESSA testing mandates. The NJDOE officials claim that the topics covered on the NJSLA in the content areas of English, math, and science are aligned to the New Jersey Student Learning Standards (NJSLS). The NJDOE states that the purpose of the NJSLA tests is “To prepare students for college and career, success in life, and work in an economy driven by information, knowledge, and innovation requires a public education system where teaching and learning are aligned with 21st-century learning outcomes” (New Jersey Department of Education, 2020, p.1). In the interpretation guide of the 2019 NJSLA, Pearson Education, the company that created the 2019 state assessment, claimed: “The primary purpose of the NJSLA is to provide high-quality assessments to measure students’ progress toward college and career readiness” (Pearson, 2019, p. 1).

In the Spring of 2019, the NJDOE first administered the NJSLA in the subject areas of English Language Arts, Math, and Science as a replacement to the Partnership for Assessment of Readiness for College and Careers (PARCC) assessment. The New Jersey Department of Education adopted the New Jersey Student Learning Standards in 2015 with only slight modifications to the Common Core State Standards that the state adopted in 2010, noting, “The state will maintain more than 80 percent of the 1,427 math and language arts standards that make up Common Core. About 230 standards will be modified” (NJSBA, 2016). Modifications did not change the substance or expectations of the standards.

Both assessments were created by Pearson; the New Jersey Education Association states: “NJSLA is a shorter version of the same PARCC test” (NJEA, n.d.). According to the NJSLA

Score Interpretation Guide from the spring of 2019, “The tests measure how well students have learned grade-level material in English Language-Arts and mathematics. Students who meet or exceed expectations are likely on track for the next grade or course and, ultimately, for college and careers” (Pearson, 2019, p. 5). A reoccurring theme with state-mandated tests administered in New Jersey is that they measure skills and knowledge that students need for college and careers.

Fact-Checking the New Jersey State Tests

Some analyses of the cognitive complexity of PARCC practice assessments have been conducted utilizing Hess’ Cognitive Rigor Matrix. Hess’ Cognitive Rigor Matrix “superimposes two widely accepted models for describing rigor- Bloom’s Taxonomy of Educational Objectives and Webb’s Depth-of-Knowledge (DOK)” (Hess et. al., 2009, p. 1). Dorrian (2021) used Hess’ Cognitive Rigor Matrix to conduct an analysis of the cognitive complexity of questions on the ELA Grade 10 and Geometry PARCC practice assessments from 2019. Dorrian’s results suggest that 93.10% of the total questions analyzed in the 2019 10th grade language arts practice test aligned with lower-level thinking (Dorrian, 2021, p. 89). The results of Dorrian's analysis of the 2019 practice questions from the geometry assessment suggested that 97.43% of the included language aligned with lower-level thinking (Dorrian, 2021, p. 90).

The low level of cognitive complexity on PARCC assessment questions extends beyond the high school grades of 9-12. Solis-Stovall (2021) found that 90% of questions analyzed from PARCC practice assessments in 3rd and 4th grade were lower-level questions, “requiring students to recall, reproduce, and use skills, and/or concepts and 10% of the questions analyzed were categorized as cognitive complex requiring strategic thinking, reasoning, and extended thinking” (p. 89). The lack of cognitive complexity calls into question the claims by education bureaucrats and policymakers in New Jersey that the state-mandated tests measure important skills necessary for college and careers.

Some of the New Jersey officials' claims may derive from similar PARCC claims about the content of its assessment. The developers of PARCC created their own definition and criteria for “high-level thinking” for mathematics and ELA frameworks and then used that definition and those criteria to code the cognitive complexity of their own test questions. Yuan and Le (2014) found, “The scoring rubric gave relatively greater weight to the difficulty of the content and relatively less weight to cognitive processes, and we found that this approach did not work well for open-ended items” (p. xii).

Sousa (2006) explained that *complexity* and *difficulty* are often used synonymously to differentiate the two terms, as complexity describes the thought process that the brain uses to deal with information. While difficulty refers to the amount of effort that the learner must expend within a level of complexity to accomplish a learning objective. Sforza et al. (2016) described complexity “as the difference between remembering a fact or imitating a procedure and developing an original product, conclusion, or process. Remembering facts and imitating procedures is less cognitively complex than developing an original conclusion, product or process” (Sforza et al. 2016, p. 9). Sforza et al. explained that difficulty is related to effort whereas complexity is related to the type of thinking required to solve a task: “Difficulty is a more static component of a learning objective that simply refers to the amount of work or effort a student must use to complete a task, regardless of complexity” (Sforza et al 2016, p. 9).

Alternative Assessments

The Council for Aid to Education (CAE, n.d.) provides a suite of assessments that claim to measure “college and career readiness skills of critical thinking, problem-solving, and effective written communication." Their assessments are preferred by some high performing private Pk-12 schools and a small group of high performing public schools around the country. CAE claims that the content and format of their tests require students to use and demonstrate problem-solving and other skills cited by the WEF, IMD and other multi-national organizations. CCRA+ claims the results of their tests can lead to a prediction of qualification of College or Career upon successful completion of High School or a Post-Secondary Institution.

Problem Statement

Although much is known about the cognitive thinking requirements, skills, and knowledge tested by current state-mandated assessments administered to millions of public school students across the country, less is known about the cadre of performance-based assessments, like the CAE tests. Assessments created by the CAE, attempt to evaluate specific types of complex thinking. The assessed skills include problem solving, analyzing and understanding data, evaluating the credibility of various documents, identifying questionable or critical assumptions, evaluating ambiguous and conflicting information, constructing organized and logically cohesive arguments, organizing and synthesizing information from several sources, and marshaling evidence from different sources in a written response. (CAE, n.d.).

School administrators are charged with preparing students to be contributing citizens to society and economically viable in an uncertain future. However, there is a gap between the skills and dispositions assessed on current state tests and the skills needed to be a contributing citizen and economically viable in an ever-changing economic environment. With standardized testing continuing as a federal mandate with the passage of ESSA, school leaders are left on their own to find ways to obtain feedback on the skills and dispositions necessary for the fourth industrial revolution.

When optimizing mandates, such as standardized testing, Tienken (2020) urges School leaders to “Gain an understanding of the reform, and then customize it at the local level to benefit students, teachers, and stakeholders” (p. 1). Finding ways to assess higher thinking is an important aspect of education reform. The WEF found that 50% of all employees will need reskilling by 2025 (WEF, 2020). Other than the state-mandated tests, some assessments claim to measure some skills recognized by global organizations as required for economic success. The CAE is the developer of one such test, as the CCRA+ claims to assess five of the skills global organizations identify as needed by their employees. Through an authentic assessment that includes document analysis and written response, more than one correct response is acceptable, as student’s critical analysis and thinking are evaluated on the CCRA+ assessment, providing real life scenarios in which the test taker applies knowledge and skills that have been acquired. There is not much independent research for school leaders to access about the kind of thinking required on the CCRA+ assessment.

Purpose of This Study

The purpose of this convergent, parallel mixed-methods study was to analyze and describe the cognitive complexity of the questions on the CCRA+, published by the CAE, compared to the language of higher-order thinking found in the research literature. CAE proclaims the CCRA+ “authentically measure the college and career readiness skills of critical thinking, problem solving, and effective written communication” (CAE, n.d., para. 2). The research questions below guide the researcher in this study.

Research Question

The study was guided by the overarching research question: What type of thinking is required by the College and Career Readiness Assessment (CCRA+) for High School Students?

Two sub-questions guided the study:

a.) What is the frequency and percentage of questions categorized as higher level within each cell of Hess' Cognitive Rigor Matrix on the CCRA+?

b.) In what ways are the skills identified as essential for success in the workplace by global organizations assessed in the CCRA+?

Research Design and Methodology

A convergent, parallel mixed-methods case study with qualitative content analysis and descriptive statistics was conducted to describe (a) the way(s) in which the language found in the questions on the CCRA+ test for high school compares with the language associated with higher-order thinking found in the research literature and (b) to describe how the skills identified as essential for success in the workplace by global organizations are assessed in the CCRA+.

Theoretical Framework

The theoretical framework for this study is guided by Dewey (1910), *How We Think*. In this seminal work, Dewey does not mention the term *cognitive complexity* and discusses *higher-order thinking* only twice; however, he describes a framework for how thought and learning progress from birth through experiences both in and outside of the school house, and through the power of reflection. Dewey (1910) wrote, "To maintain the state of doubt and to carry on systematic and protracted inquiry– these are the essentials of thinking" (p. 15). Dewey (1910) defined reflective thinking as the "Active, persistent, and careful consideration of any belief or supposed form of knowledge in the light of the grounds that support it, and the further conclusions to which it tends, constitutes reflective thought" (pp. 5-6).

Dewey (1910) warned of the negative consequences of schools, focusing only on lower-level thinking at the expense of higher-order thinking, especially reflective thinking. He wrote that lower-thinking was not the aim of schools, “While it is not the business of education to prove every statement made, any more than to teach every possible item of information” (p. 32). Refocusing the purpose of education to be a place where students learn how to think was of importance to Dewey:

> It is its (schools) business to cultivate deep-seated and effective habits of discriminating tested beliefs from mere assertions, guesses, and opinions; to develop a lively sincere, and open-minded preference for conclusions that are properly grounded, and to ingrain into the individual’s working habits methods of inquiry and reasoning appropriate to the various problems that present themselves. (p. 30)

Hess’ Cognitive Rigor Matrix served as the framework for this study to conceptualize Dewey’s higher-order thinking ideas. Hess’ Cognitive Rigor Matrix integrates Bloom’s Revised Taxonomy and Webb’s Depth of Knowledge in a matrix that can be used to categorize the complexity and type of thinking required by education standards, learning tasks, and assessment items.

Limitations of the Study

For the purpose of this study, cognitive complexity is determined from the categories of Hess’ Cognitive Rigor Matrix. While Hess’ Cognitive Rigor Matrix was published in 2009, limited research is available in comparison to other matrices and frameworks used to determine cognitive complexity.

The research was conducted using a case study design and was limited to the analysis of retired, publicly available middle school and high school assessment prompts of the CCRA+ that are no longer in use. Analysis cannot determine if actual test items would demonstrate similar cognitive complexity as the assessment continues to evolve. Due to the limited number of practice tests available, a small number of questions were available for analysis. The case study design limits the generalizability of the results beyond the grade levels and the editions of the CCRA+ examined for this study.

Delimitations

Delimitations of this study include the assessment that was analyzed, the CCRA+. Different assessments could gauge College and Career Readiness for students through asking questions of varying cognitive complexity. Only retired, publicly available assessment prompts for Grades 6-8 and Grades 9-12 were available for analysis. The analysis of a limited availability of assessment prompts cannot guarantee the findings of cognitive complexity are reliable for other versions of the same assessment. The publicly available assessment prompts have been retired and are no longer in use, this also could result in a different level of thinking than what is assessed on current versions of the exam. Hess' Cognitive Rigor Matrix for English Language Arts was the tool utilized to determine cognitive complexity. A different framework could lead to different results of cognitive complexity per question.

Definition of Terms

Bloom's Revised Taxonomy: Northern Illinois University's Center for Innovative Teaching and Learning writes "Bloom's Revised Taxonomy is one of many tools that faculty can use to create effective and meaningful instruction. Use it to plan new or revise existing curricula; test the relevance of course goals and objectives; design instruction, assignments, and activities; and develop authentic assessments" (2020).

Cognitive Complexity: The state or quality of a thought process that involves numerous constructs, with many interrelationships among them. Such processing is often experienced as difficult or effortful (American Psychological Association, 2022).

Critical Thinking: While there is not one definition of critical thinking, Ennis (2016) defines such thinking as "reasonable reflective thinking focused on deciding what to believe or do" (Ennis, 2016). Scriven and Paul (1987) define critical thinking as "the intellectually disciplined process of actively and skillfully conceptualizing, applying, analyzing, synthesizing, and/or evaluation information gathered from, or generated by, observation, experience, reflection, reasoning, or communication as a guide to belief and action" (1987).

Cut Score: Selected points on the score scale of a test which is used to determine whether a particular test score is sufficient for some purpose (Educational Testing Services, 2006).

Functional Fixedness: The inability to realize that something known to have a particular use may also be used to perform other functions. When one is faced with a new problem, functional fixedness blocks one's ability to use old tools in novel ways. (Britannica, n.d.)

Global Competitiveness: The set of institutions, policies and factors that determine the level of productivity of a country. A country is competitive based on how it promotes the well-being of its members (World Economic Forum, 2016).

Hess' Cognitive Rigor Matrix: Hess, et al. (2009) define Hess' Cognitive Rigor Matrix as superimposing two widely accepted models for describing rigor – Bloom's Taxonomy of Educational Objectives and Webb's Depth-of-Knowledge model for analyzing instruction and enhancing lesson planning (Hess, et al. 2009).

Higher-order thinking: Skills that go beyond basic observation of facts and memorization. Higher-order thinking skills are used when students are evaluative, creative, and innovative (University of Connecticut Center for Excellence in Teaching and Learning, n.d., para. 1).

Reflective Thinking: active, persistent and careful consideration of any belief or supposed form of knowledge in the light of the grounds that support it, and the further conclusions to which it tends (Dewey 1910).

Selected-Response Questions: Selected-response items are those in which students read a question and are presented with a set of responses from which they choose the best answer (National Assessment Governing Board, 2023). Selected-response questions are also referred to as multiple choice questions.

Standardized Assessment: An assessment instrument whose validity and reliability have been established by thorough empirical investigation and analysis. It has clearly defined norms, such that a person's score is an indication of how well he or she did in comparison to a large group of individuals representative of the population for which the test is intended. An assessment instrument administered in a predetermined manner, such that the questions, conditions of administration, scoring, and interpretation of responses are consistent from one occasion to another (American Psychological Association, 2022).

Student Achievement: Definitions of student achievement vary dependent on educational philosophy. The education company Top Hap provides two definitions for student achievement: Student achievement is the measurement of the amount of academic content a student learns in a given time frame. Each instruction level has specific standards or goals that educators must teach to their students. Achievement is usually assessed through frequent progress and comprehension checks and examinations. There is no consensus on how it is best evaluated or which elements of it are most important. (Top Hat, n.d., para. 1) The second provided definition from Top Hat is: Student achievement refers to the extent to which a learner has attained their short or long-term educational goals. Individual differences in academic performance are strongly correlated with differences in personality and intelligence. As well, students' levels of self-efficacy, self-control and motivation also impact levels of achievement (Top Hat, n.d., para. 2).

Webb's Depth of Knowledge: A framework that designates how deeply students must know, understand, and be aware of what they are learning in order to attain and explain answers, outcomes, results, and solutions (Francis, 2018, para. 12).

Organization of Dissertation

The dissertation was organized to introduce the problem presented in terms of the claim of cognitive complexity on student assessment and public schools' preparation in developing economically viable citizens. In Chapter II the literature review included the theoretical framework of the study, the previous research and findings pertaining to the cognitive complexity of questions, higher-order thinking, the cognitive complexity of current standardized assessments, the skills identified as essential for employees, and the current status of student preparation of mastering these skills in the United States. In addition to the findings, Chapter II included a synthesis of the research as well as a discussion of areas that have not yet been analyzed. In Chapter III, the overview, method of the study, and procedures were explained. Chapter IV provided the findings of the study inclusive of the data found. In closing, Chapter V summarized the findings, discussed the implications for students, and provided recommendations for future research and schools in terms of procedures, practices, and policy formation.

Chapter II

Literature Review

The literature review served four purposes. The *first* purpose was to identify empirical studies and findings on higher-order thinking and cognitive complexity. The *second* purpose of the literature review was to analyze relevant theories and constructed definitions of higher-order thinking and cognitive complexity, such as Bloom's Taxonomy, Webb's DOK, and Hess' Cognitive Rigor Matrix. The *third* purpose was to review the history of standardized testing in the United States from the lens of legislation and policy. *Finally*, literature was reviewed comparing the higher-ordering thinking and cognitive complexity of questions on standardized testing, inclusive of CCRA+, PARCC, and other commonly required for high school graduation or college admission assessments. The literature review concludes with the theoretical framework of this dissertation, a focus on experimentalism and progressivism in education through the work of John Dewey in *How We Think* (1910) and *The Child and the Curriculum* (1902).

Literature Search Procedures

This review follows guidelines for a scholarly literature review developed by Boote and Beile (2005). The researcher's review of the literature included online searches from the Seton Hall University Database, New Jersey Department of Education websites, and previous dissertations, pertinent legislation and policies, and peer-reviewed articles. Searches included, but were not limited, to the keywords higher-order thinking, cognitive complexity, standardized testing, cut score, Webb's Depth of Knowledge, Bloom's Taxonomy, and Hess' Cognitive Rigor Matrix. Definitions for the Key Terms found in Chapter I were found from the journal articles found following these online searches, textbooks, and seminal writings.

Criteria for Inclusion of Literature

The research used in this review included:

A.) Dissertations on higher-order thinking

B.) Peer-reviewed and non-peer-reviewed articles focused on higher-order thinking, cognitive complexity, and problem-solving.

C.) Government and legislative reports

D.) Professional Development books for education

E.) Books on the topic of thinking

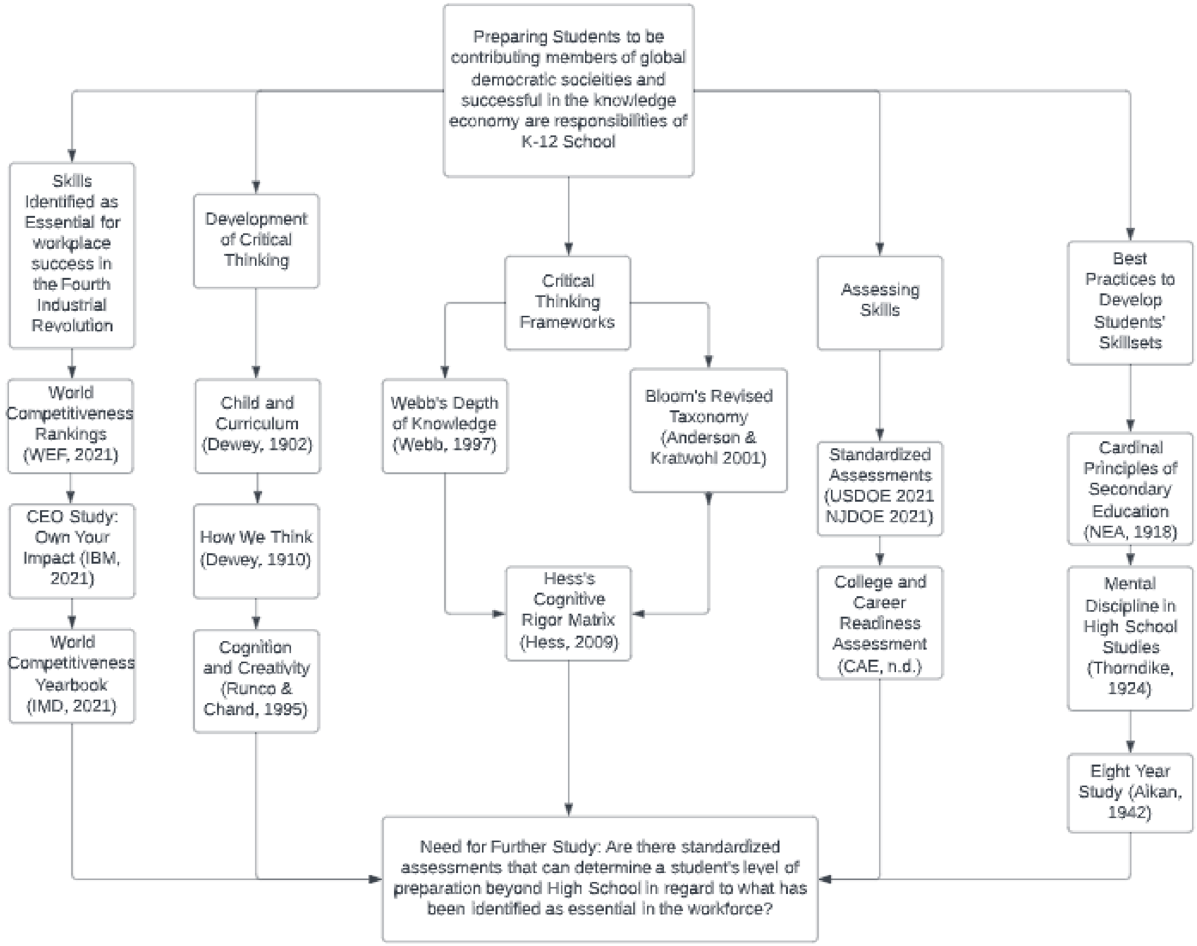
F.) Seminal works on the development of thinking

Literature Map

The Literature Map represented in Figure 1 provides a visual of the paramount publications, questions, and focus of the study.

Figure 1

A Literature Map: An Alignment with Research Questions



Note. The Literature Map above depicts the prominent literature and focus of the study.

Methodological Issues with Existing Studies

The first methodological issue on the topic of cognitive complexity and higher-order thinking is that an exact definition is challenging to find and not consistently agreed upon. Although both phrases are found frequently in articles written by educational associations and organizations selling products to increase the skills of cognitive complexity and higher-order thinking, an exact definition is elusive. Moore (2011), writes, “there is broad agreement about the importance of critical thinking as an educational ideal, a view often expressed in the literature is that academics are not always so clear about what the concept means" (p. 506). Without a specific definition of cognitive complexity or higher-order thinking, educators rely on a number of frameworks to determine the level of cognitive complexity of higher-order thinking. Additionally, the terms cognitive complexity and higher-order thinking are often used interchangeably, as well as the skills comprising of these two terms, such as creativity, problem solving, and critical thinking to name a few. Non-peer reviewed articles give examples of lessons and projects yet do not give a formal definition of cognitive complexity or higher-order thinking. Mulnix (2012), concludes that a "difficulty with determining whether critical thinking can be taught, or even measured, is that there is widespread disagreement as to what critical thinking actually is or amounts to" (p. 464). Finally, the term 21st century Skills are often used synonymously with both higher-order thinking, which is also an ambiguous term in education. When writing about 21st century skills, Silva, (2009) conveys, "it's no wonder that the term seems vague or confusing. There are hundreds of descriptors of the skill set, including life skills, workforce skills, interpersonal skills, applied skills, and noncognitive skills" (p. 630).

Although there is much literature on standardized testing, few studies focus on the CCRA+ or the previous version of the same assessment, the College and Work Readiness Assessment Plus (CWRA+). Additionally, the data that does come from the assessments provided by the CAE are not representative of an extensive or diverse population. The majority of institutions using this assessment are post-secondary or Grade 6-12 private independent schools. The lack of research leads to claims of higher-order thinking skills being critiqued via the authentic assessment on the CCRA+ without much data supporting if the statements are accurate and can be determined from the assessment.

Review of Literature Topics

Fourth Industrial Revolution's Impact on Education

The Fourth Industrial Revolution presents a need for K-12 schools to focus on ensuring students are provided with the skills essential to success in the labor market. The shift in this revolution is the enhancement of Artificial Intelligence (AI), robotics, and data which will ultimately replace jobs leading to unemployment and a widening global social economic gap. If the skills and abilities developed through Grade 12 can be replaced by a robot or computer program that can complete the work of individuals who have not created the capacity to think critically, innovatively, or flexibly schools are not preparing students for college or a career. Dell Technologies *Future of Work Report* (2019) described a workplace that will include humans and machines completing tasks simultaneously. The ability for humans to contribute will depend on higher-order thinking and emotional intelligence (pp. 10-11). For this reason, in order to maintain employment, critical thinking must be developed throughout students' time in primary and secondary school.

The Fourth Industrial Revolution provides an opportunity for schools to focus on the essential needs for individuals to successful on a global platform. In order for students to understand their role beyond the K-12 education sector, Kayembe and Nel (2019) wrote "students who study the humanities and social sciences need to understand at least the foundation on which AI is based and operated…the Fourth Industrial Revolution drives the idea of a multidisciplinary field, whereby humanities and social sciences join technologies to solve problems (p. 87). The notion of teaching students to problem solve is the essence of critical thinking. The Institution of Engineering and Technology (IET) identified "What industry needs from the education system is a diverse talent pool, young people who can solve problems and who can readily apply maths, computing, science and humanities knowledge, and know-how to deliver societal and technological benefit" (2018, p.2). For students to be prepared to contribute beyond secondary school, they must have developed "the ability to think, to learn, to solve problems, to be inquisitive, entrepreneurial, proactive and creative" (IET, 2018, p. 2).

Higher-Order Thinking and Cognitive Complexity

Critical Thinking

A skill identified as essential in K-12 education, Higher Education, and the workplace, critical thinking does not have an agreed upon definition. Hitchcock (2018) reports that Ennis (2016) has found 17 definitions for critical thinking, 14 of which are philosophically oriented scholarly definitions and three are dictionary definitions. Ennis (2016) defines critical thinking as "reasonable reflective thinking focused on deciding what to believe or do" (p. 166). Ennis (2016) cites the critical thinking definition by Scriven and Paul (1987) as an alternate description of the skill "Critical thinking is the intellectually disciplined process of actively and skillfully conceptualizing, applying, analyzing, synthesizing, and/or evaluation information gathers from, or generated by, observation, experience, reflection, reasoning, or communication, as a guide to belief and action" (p. 166). Many of these skills are utilized in the higher-order thinking Frameworks to determine the cognitive complexity of the task or question being assessed. Furthermore, the use of reflective thinking defined by Dewey (1910) as "Active, persistent, and careful consideration of any belief or supposed form of knowledge in the light of the ground that support it and further conclusions to which it tends, constitutes reflective thought" (p. 5). In order to develop a thought process that can be defended and replicated, reflective thinking is a requirement. The acceptance of facts and information without thinking into the reasoning or process in which an outcome is created provides examples of lower-level thinking that would not be considered critical thinking.

Higher-Order Thinking Frameworks

Bloom's Taxonomy of Educational Objectives (Bloom et al., 1956) and Webb's Depth of Knowledge (Webb et al., 2005) are some examples of frameworks used in K-12 schools to focus curricula writing, lesson planning, and assessment creation. Although the two frameworks focus on cognitive complexity, "Bloom's taxonomy categorizes skills required of the brain when faced with a new task, therefore describing the type of thinking processes necessary to answer a question" (Hess, 2009, p. 4), whereas Webb's Depth of Knowledge "relates more closely to the depth of content understanding and scope of a learning activity which manifests in the skills required to complete the task from inception to finale" (Hess, 2009, p. 4). Hess' Cognitive Rigor Matrix, overlays Bloom's Revised Taxonomy on Webb's Depth of Knowledge thereby combining the cognitive process and depth of learning in tasks for a straightforward mechanism to determine the level of thinking required to successfully respond to questions and demonstrate understanding. Knowledge is the foundational level of Bloom's original Taxonomy, correspondingly, recall captures the first level of Webb's Depth of Knowledge.

Difficulty versus Complexity

Tienken (2017) writes "There seems to be some confusion on the part of proponents of standardization about the difference between difficulty and complexity as related to high-level thinking" (p. 122). Tienken (2017) cites Sousa (2011) as "explained complexity as the thought processes required to address a given task, a set of tasks, or a problem" (p. 123). Sforza, et al. (2016) defined "difficulty is a more static component of a learning objective that simply refers to the amount of work or effort a student must use to complete a task, regardless of complexity" (p.

9). Objectives, questions, and assessments can be more difficult without requiring a higher-order of complexity. When “one objective simply requires more steps, and thus more effort” (Tienken, 2017, p. 123), the assessment is not necessarily requiring more or higher-order thinking, but potentially just spending more time on coming to a solution.

Although previous research has shown standardized tests have asked difficult, not complex questions, the differentiation between the two types of questions has been noted. The Mississippi Department of Education (MDE) includes in their description of Webb’s Depth of Knowledge (1997), “The Depth of Knowledge (DOK) level should reflect the complexity of the cognitive processes demanded by the task outlined by the objective, rather than its difficulty. Ultimately the DOK level describes the kind of thinking required by a task, not whether or not the task is “difficult” (p.5) Tienken (2017) writes “Making students perform mental acrobatics is not complex…That is wasting time and money on contrived difficulty.” (p.123) Francis (2014) adds to the difference of complexity and difficulty saying “questions, problems, or tasks are defined as easy or hard and are determined by how many people can answer the question, address the problem, or accomplish the task correctly or successfully (para. 3). In comparison, Francis (2014) writes “complex questions, problems, and tasks also allow students to delve deeper into the content, concepts, ideas, subjects, and topics being taught” (para. 7) Sforza, et al. (2016) write “Although complexity and difficulty are necessary components of an intended curriculum, the Depth of Knowledge or complexity of a learning objective is dynamic and encompasses the multiples dimensions of an objective. (pp. 9-10). Making questions and problems more difficult and time-consuming does not mean students are solving complex problems, thinking critically, or transferring knowledge.

Standardized Tests

Standardized tests are a focus of education policy, with their results influencing a multitude of educational decisions. Standardized testing has emerged as the preferred policymaking tool to hold schools accountable for teaching the state-mandated curriculum standards. The results from state standardized tests are often used as the main factor for declaring students prepared for college and careers. A single-cut score on an assessment is the usual deciding factor on whether students are college or career ready. Schools and teachers are then evaluated based on having met this target, based on the percentage of an entire student body, not individual student performance. Bracey (2008, para. 1), defined cut scores as "a score that determines whether a student passes or fails, is proficient or not or is being educated or left behind." Consistency has not been established when determining cut scores, depending on the organization that created the assessment different methods are utilized to select a cut score indicating proficiency.

For example, "The ACT College Readiness Benchmarks are scored on the ACT subject area tests that represent the level of achievement required for students to have a 50% chance of obtaining a B or higher or about a 75% chance of obtaining a C or higher in corresponding credit-bearing first-year college courses" (ACT, 2019, p. 2). Educational Testing Service (ETS) outlines a multi-step process to determine cut scores. Zieky and Perie (2006, p. 7) described different processes for cut score setting procedures that could require judges to estimate the probability that a hypothetical group of students knows the answer to a test question, better known as The Rasch Rating Scale Model. Another type of procedure requires judges to examine a student's performance and to decide whether the performance is good enough for some particular purpose. Zieky and Perie, close with, "No purely objective method exists. There are no "true" cut scores that a group of perfectly selected, perfectly trained judges using a perfect method will find" (2006, p.7). Without consistent practice, process, or definition of what a cut score is students and educators are left confused by the inconsistencies found when determining what is considered proficient based on a test score.

New Jersey Legislation informing Standardized Testing

Beginning in 1975, the NJDOE required standards-based assessment, which has been referred to as standardized testing throughout this study. During this approximately 50-year time frame, a cycle has emerged of legislation passing, a new test being adopted, a trial period that allowed for districts to align curricula to teach to the test, the test becoming a graduation requirement, new standards being adopted, with a claim of providing students with "educational opportunity" grounded in "rigorous state content standards" (New Jersey Department of Education, 2016, para. 1, 6).

The Public School Education Act (PSEA) passed in 1975, was adjusted in 1976 to meet the requirements of being the graduation assessment requirement for students, who then took the Minimum Basic Skills (MBS) assessment in third, sixth, and ninth grade. Passing the MBS officially became a graduation requirement for students in the graduating class of 1982.

In 1983, the NJDOE adopted the High School Proficiency Test (HSTP9) "a more challenging assessment to measure the minimum skills in reading, writing, and mathematics." (New Jersey Department of Education, 2016, para. 3). This assessment then changed names from HSPT9 to HSPT 11, in 1988, following legislation that moved the assessment from ninth to eleventh grade. This newly adopted legislation included the addition of an eighth-grade standardized assessment known as the Early Warning Test (EWT). Following true to the cyclical nature of adopted standardized assessments, the HSPT11 because the graduation requirement for all students in 1993 after two years of pilot testing.

May of 1996, brought a significant change to the creation of standardized assessments in the New Jersey Public Schools. The change was sparked when the NJDOE "adopted the Core Curriculum Content Standards (CCCS) which enumerated what all New Jersey students should know and be able to do by the end of the fourth and eighth grades, and upon completion of a New Jersey public education" (NJDOE, 2016, para. 5) The newly adopted standards lead to three new standardized assessments, the Elementary School Proficiency Assessment (ESPA), the Grade Eight Proficiency Assessment (GEPA), and the High School Proficiency Assessment (HSPA). These assessments were administered from 1997 through 2008, depending on the assessment.

Following the No Child Left Behind Act of 2001 (NCLB), standardized assessment in New Jersey once again changed. The federal legislation now required yearly state testing in math and English Language Arts (ELA) for Grades 3 through 8 and once in High School. In response to NCLB, the NJDOE added in third grade The New Jersey Assessment of Skills and Knowledge (NJ ASK) in 2003 as well as a fourth grade science assessment. Public school students in the state of New Jersey now were administered the NJ ASK in Grades 3-8 in math and ELA, HSPA, and the Alternate Proficiency Assessment (APA) "for students with severe cognitive disabilities, and end of course assessments in Biology and Algebra" (New Jersey Department of Education, 2016, para. 7).

New Jersey revised statute Title 18A:&C-1 - Commissioner of education to develop a program of standards and guidelines in the year of 2013. Prior to the revised statute, the State of New Jersey required "By July 1, 1980, the Commissioner of Education with the approval of the state Board of Education shall establish a program of standards for graduation from secondary school" (Justia, 2013, para. 1).

In the 2014-2015 school year, state testing was administered electronically for the first time when the PARCC assessments became the required test for grades 3 through 11 in math and ELA. Testing in science was required once in grades 3 through 5, grades 6 through 8, and High School and remained in the paper and pencil format. In 2016, the NJDOE adopted the New Jersey Student Learning Standards (NJSLS), which lead to a change in the name of the state standardized test to the New Jersey Student Learning Assessment. At this time, the science assessment also moved to an electronic format.

Standardized testing was canceled in both the 2019-2020 and 2020-2021 academic school years due to the Coronavirus disease 2019 (Covid-19) pandemic. The deferral of testing in the spring of 2021 led to an additional standardized test in the fall of the 2021-2022 school year, to "satisfy the federal statewide assessment requirement to administer general assessments". (New Jersey Department of Education, 2021, para. 1). With this change, the state of New Jersey required the New Jersey Graduation Proficiency Assessment (NJGPA) for the first time in March of 2021, as the graduation assessment for all students in the graduation class of 2023 and beyond.

CCRA +

The Council for Aid to Education (CAE) has created a suite of assessments with the claim of being able to "situate students in real-world, complex decision environments where they must analyze and synthesize data, address important issues, propose solutions, and recommend courses of action to resolve conflicts" (CAE, n.d., para. 7). In 2007, CAE partnered with the Rand Corporation to create the College and Work Readiness Assessment Plus (CWRA+) for students in grades 6-8 and 9-12 (Rand, 2018). In 2019, CAE began providing the CWRA+

independently of the Rand Corporation with the retitled name of the College and Career Readiness Assessment (CCRA+). CAE published the belief that "Traditional multiple-choice assessments miss the mark on providing data about the foundational skills that are essential to students' academic and career success - critical thinking, problem-solving, and written communication (CAE, n.d., para. 1)"

The skills the CAE attempts to assess through the CCRA+ align with what the WEF has found essential to be prepared for success. Furthermore, these skills are what John Dewey (1910) identified as the components of learning and thought. After being awarded a 2021 Tech Edvocate Award Finalist for Best Learning Assessment or Tool (Global News Wire, 2021), Yayac, the President and CEO of CAE stated "While content knowledge is important, it is insufficient to prepare the next generation of students for their next step" (para. 4). The ability to transfer knowledge is a necessary skill, which is often not met without reflection or thought or when students are trained in isolated exercises (Dewey, 1910).

Theoretical Framework

Theoretical frameworks to guide the study of higher-order thinking are varied. This study utilized Hess' cognitive rigor matrix as a framework to define higher-order thinking and as a tool to determine the level of rigor, deemed as required to "succeed in future college-entry courses and workforce training programs" (Hess, et al, 2009, p. 1). As one reviews work from John Dewey, landmark works such as *The Cardinal Principles of Secondary Education* and the *Eight-year Study*, the skills identified as essential in the early 20th century in comparison to the skills needed as reported by leading organizations within the past century are similar. Silva (2009, p. 631), in an article titled Measuring skills for 21st century learning noted, the century-specific label is also misleading. Knowing how to think critically, analytically, and creatively are not skills specific or unique to the 21st century, much of the same has been argued by philosophers and educators from Socrates to 20th century John Dewey. These skills have been written about and easily accessible to educators, yet are still seldom the focus of curricula, units, and lessons.

Dewey and a Framework for Thinking

Dewey's progressive educational philosophy stood on the foundation that learners are active participants in which students learned best by asking a question, creating a process to test, analyzing results, and reflecting on what was found. Similar to the scientific method, Dewey believed students formed lasting knowledge through experimentation. "Active, persistent, and careful consideration of any belief or supposed form of knowledge in the light of the grounds that support it, and the further conclusions to which it tends, constitutes reflective thought" (Dewey, 1910, pp. 5-6).

Dewey (1910) suggested that schools should ensure that students are exposed to opportunities that teach them and allow them to think at high levels instead of simply training students to be receptacles of information:

> It is its (schools) business to cultivate deep-seated and effective habits of discriminating tested beliefs from mere assertions, guesses, and opinions; to develop a lively sincere, and open-minded preference for conclusions that are properly grounded, and to ingrain into the individual's working habits methods of inquiry and reasoning appropriate to the various problems that present themselves. (p. 32)

Furthermore, Dewey (1910) warned of the misconception of associating gathering of all possible knowledge and facts with learning, "learning is not wisdom; information does not guarantee good judgment" (p. 127). Dewey (1910, p. 71) warns throughout *How We Think* of the constraints put on student thinking, curricula, and assessment when standardization and testing are put into practice in schools. To maximize the evolution of student thinking, each student progression must be embraced. As an aid in the development of student thinking, the learner must have the ability to be curious, create inferences, and experiment. Spontaneous play naturally creates problem-solvers through the transfer of knowledge and observation. Therefore, it is essential for students to have the opportunity to explore and transfer knowledge and skills acquired in one setting to another.

Dewey's ideas about thinking (1910) are now echoed by CEOs of multinational companies of the need for more critical thinking, and creativity in the workplace (WEF, 2020) In order to develop thought and make connections, students must utilize time to reflect on the instruction. This is where students can build upon a lesson or creatively connect and expand upon previous topics and leaves them with the ability to use their imagination to create new meaning. Dewey (1910) recommends, "Let the facts be presented so as to stimulate imagination, and culture ensures naturally enough" (Dewey, 1910, p. 270). In addition to the capability to think critically, through the process of thought identified by Dewey, the learner also has the opportunity to be creative when allowing their imagination to apply the knowledge and facts they have required transferring information from a lesson previously learned to the new material. This creativity poses the opportunity for innovation, and complex problem solving as students are given boundaries to work within but move beyond what is often seen when preparing for tasks and answering questions, which align with the cells of the Cognitive Rigor Matrix, indicating the lowest level of thinking. Dewey (1910) warns of "Overdoing the mechanical and automatic "Drill". The danger in those studies where the main emphasis is on skill acquisition is just the reverse. The tendency is to take the shortest cuts possible to gain the required end (p. 58). When focusing on knowledge acquisition, remembering facts, or being able to recall, students suffer as "This makes the subjects mechanical, and thus restrictive of intellectual power" (Dewey, 1910, p. 58). Students are not learning how to think at higher-order levels when their education is situated around recall.

A responsibility of teachers is to guide and model when necessary, exposing students to practices that will lead to their cognitive development. Dewey (1902) writes, "The child is expected to "develop" this or that fact or truth out of his own mind. He is told to think things out, or work things out for himself, without being supplied any of the environmental conditions which are requisite to start and guide thought" (p. 16). The lower levels of thinking, as defined by Bloom or Webb do not take any skill or process from the teacher, as a student who demonstrates a desire can memorize, without putting meaning or thought behind the concept, words, or skillset.

Recall is passive learning, as it does not draw on experience or higher-order levels of thinking. As Dewey (1902) provides a further example of the irrelevance of needless memorizing compared to experiences when giving an example of "the difference between the notes which an explorer makes in a new country, blazing a trail and finding his way along as best he may, and the finished map that is constructed after the country has been thoroughly explored" (p. 17). Two people can claim they learn of the land through two different methods, the student in a classroom views a map, the explorer connects through personal experience. From here, Dewey (1902) differentiates writing "Development does not mean just getting something out of the mind. It is a development of experience and into an experience that is really wanted. (p. 16)

The development of the mind and gaining knowledge comes through the process of individual experiences, the introduction of knowledge, reflection and the connections the student can make between these events. “Any significant problem involves conditions that for the moment contradict each other. Solution comes only by getting away from the meaning of terms that is already fixed upon and coming to see the conditions from another point of view (1902, p. 5). Through the process of knowledge development, a student are presented with new information, knowledge, or experiences that conflict with what they have previously known. Dewey extends, writing, “this reconstruction means travail of thought. Easier than thinking with surrender of already formed ideas and detachment from facts already learned is just to stick by what is already said, looking about for something with which to buttress it” (1902, p. 5). The development of critical thinking is the development of student motivation to actively pursue knowledge and conflict between previously established beliefs. Memorizing and recalling information is what Dewey refers to as the “easier” way of thinking, as ideas have already been formed, many times by the teacher and taught to the student with a lack of connection or background understanding. Dewey (1902) writes, “it is easier to see the conditions in their separateness, to insist upon one at the expense of the other, to make antagonists of them than to discover a reality to which each belongs” (pp. 5-6).

Disengagement of thought and learning develops when education experiences do not foster students making connections between their personal experiences and the content. Students must be allowed to grapple with information in ways that build upon previous knowledge and experiences. Dewey writes “the lack of any organic connection with what the child has already seen and felt and loved makes the materials purely formal and symbolic” (1902, p. 20). Providing students with maps, figures, diagrams, or video clips as seen on current standardized testing does not provide a lasting impact or connection for students if there is not an experience to reflect or build upon. When a connection lacks, students will lack the motivation to learn or research new information on the topic (Dewey p. 21). When students are provided questions and assessments that do not require critical thinking, Dewey apprises “the child’s reasoning powers, the faculty of abstraction, and generalization, are not adequately developed” (1902, p. 22). Dewey (1902) closes *The Child and the Curriculum* with an analogy that calls out the concept of difficult “sugar-coated” in lieu of complex thinking “Mental assimilation is a matter of consciousness; and if the attention has not been playing upon the actual material, that has not been apprehended, nor worked into faculty” (p. 26) Both the *Child and the Curriculum and How We Think* provide a framework to develop questions, lessons, and assessments that ultimately develop a critical level of thinking at a cognitively complex level.

Following in the footsteps of *How We Think* and *The Child and the Curriculum,* The Commission on the Reorganization of Secondary Education presented the *Cardinal Principles of Secondary Education* was published in 1918 following a report to the National Education Association in 1911 with a focus on high school’s curriculum and how it prepared students for both vocation and post-secondary institutions. The reviewing commission, 1918, wrote “Secondary education should be determined by the needs of the society to be served” (p. 7). The responsibility of the school is to prepare students who are ready and able to contribute to society upon completion of school. The first factor noted as a reason for the reorganization of secondary education is “Changes in society” (NEA, 1918, p. 7). As we have embarked on the Fourth Industrial Revolution, the thinking requirements of assignments and the structure of school have not adjusted to what is being called on of our graduating seniors. The Commission (1918) also noted the importance of students applying the knowledge they have learned, when writing, “teaching methods must be tested in terms of the laws of learning and the application of knowledge to the activities of life, rather than primarily in terms of the demands of any subject” (p. 8). One of the foundational principles of secondary education was a vocation. In order to “secure a livelihood for himself and those dependent on him, to serve society well through his vocation” (NEA, 1918, p. 13) Individuals must be prepared with higher-order thinking abilities to be successful in their field while contributing to society.

Among the activities noted by leading organizations as essential, while including higher levels of cognitive complexity that were suggested by the commission (1918) include “the assignment of projects and problems to groups of pupils for a cooperative solution and socialized recitation” (p. 14). The commission (1918) did not use the term critical thinking. They implied the need for complex thinking when the authors of the report wrote, “Both of these devices give training in collective thinking” (p. 14). Creativity and problem solving are at the heart of solving authentic problems and creating projects, identified as requisites for the worker of the 21st century and beyond. The commission (1918) continued to describe their observation of passive learning in core areas when writing the observation “Too frequently, however, does mere information, conventional in value and remote in its bearing, make up the content of social studies” (p. 14) This form of learning, would correlate to the lowest level of recall or memorization which does not provide students the opportunity to make connections or think critically.

Echoing the progressive philosophy of Dewey and following the publication of *The Cardinal Principles of Secondary Education*, the Progressive Education Association (PEA) spearheaded the Eight-Year Study that included thirty high schools from 1933 to 1941. When comparing these two seminal works, one will find a synergy of the “nine areas of human activity and problems of living as the basis of curriculum construction" (Aikin, 1942, p. 74) and the seven objectives of the Cardinal principles. The work of the 30 schools emanated from an overarching question of "How can the high school improve its service to American Youth?" (Aikin, 1942, p. 1). The PEA (1942) reported, "whatever employment conditions are at any time, the school admits the inescapable responsibility of helping … to prepare for economic self-support and useful service to the community" (p. 66). The PEA (Aikin, 1942) continued, "Few schools anywhere have met this responsibility fully" (p. 66). This sentiment was already stated from the viewpoint of various organizations in their CEO reports. Of the ten objectives created by the participating schools "The development of effective methods of thinking" (Aikin, 1942 p. 89) was the first listed.

The feedback from high school employees was that the ideas proposed would benefit students in their development but at the potential harm to college admission. This concern led to thirty high schools being a part of a project organized by the PEA. The findings from educators in the 1930s are parallel to the areas of need for public schools in our current day, "Our secondary schools did not prepare adequately for the responsibilities of community life. Schools generally were excellent examples of autocratic rather than democratic organization and living (Aikin, 1942, p. 4). Continuing on with even greater caution, the PEA found:

> The high school seldom challenged the student of first-rate ability to work up to the level of his intellectual power. It was easy for him to "get his lessons", pass his courses. The result was that many a brilliant mind developed habits of laziness, carelessness, superficiality. These habits become firmly established during adolescence, prevented the full development of powers. Even the conscientious student of superior ability did not find himself in seriously involved in a great intellectual enterprise. Seldom was any

> student "set on fire" intellectually, ready to explore on his own, ready to conquer difficulties. (p. 5)

Noted through these observations, students did not express complex levels of cognitive thinking, creativity, or problem solving. Lessons were focused on turning around a product without much thought or engagement. Observation continued by the PEA (1942) "In spite of greater understanding of the ways in which human beings learn, teachers, persisted in the discredited practice of assigning tasks meaningless to most pupils and of listening to recitations. The work was all laid out to be done" (p. 5). Students were not provided the opportunity to introduce questions that piqued their interest, subjects were taught in isolation, recall and recitation were the basis of classroom discussion and assessment. Through low levels of cognitive complexity in the classroom and cross-curricular lessons, units, or projects, students struggled to build connections in understanding how the lessons could make meaning to their lives.

As the PEA looked into curricula changes needed, suggestions focused on students being provided the opportunity of authentic learning and problem solving that allowed students to apply the knowledge from one content area across subjects and to areas of interest for the students. For example, the PEA (1942) declared the thirty schools participating in the study observed current events and applicable information to the daily life of the students:

> the science class might be studying the technique of solving problems, not only in the field of science but in many other phases of life. The class in Spanish might be investigating the influence of geography upon the life and character of South American peoples. The group in mathematics might be applying principles of logic to an analysis of

> a local problem of housing or conservation. The class in history might be drawing up a statement for the next school assembly, outlining the issues involved in the annual election of student leaders. The English class might be analyzing recent newspapers and magazines to discover ways and means by which propaganda molds public opinion. (p. 46, 47)

Some schools participating in the study also required an internship, for upperclassmen which could be reported in a variety of ways, inclusive of "personal investigation, interviews, and work with one's hands (Aikin, 1942, p. 48). This form of project provides students the opportunity to learn a field or topic of interest at the highest levels of cognitive complexity. The science classes found that "The immediate purpose is satisfaction of the pupils' desire to know and understand; but the larger purpose may be to develop habits of critical thinking and intellectual honesty" (Aikin, 1942, p. 50). Through the development of critical thinking, according to some of the participating schools (Aikin, 1942) the goal of the three year science course sequence in the participating schools was

> A willingness to experiment and to accept the conclusions reached from experiments; a critical attitude toward authorities; an attitude of suspended judgment; recognition that all theories are tentative and all truth relative; an awareness of the possibilities open to man through his understanding of the laws of life. (p. 51, 52)

Several of the high schools that participated in the study included a senior internship in which students worked at businesses in the community. Through these experiences, students were able to apply what was previously taught in subject silos. As Aikin described (1942), the opportunities "become related sources of knowledge and understanding as they contribute to the student's purposes of making a living and doing useful work in which he finds growth and satisfaction" (p. 57). Including a senior internship allowed both companies and students to gauge if they had developed the skills needed to be successful in the workplace upon high school. Importantly, participation in an internship provides students with the ability to have experiences to connect to the curricula as was deemed critical by Dewey in 1902. Assuming the feedback given by the cooperating organizations was honest, direct, and aligned with the skills deemed essential, students were given the opportunity to understand if they were prepared to be successful and if not had time to learn the necessary skills from their teachers prior to graduation. Students who participated in two-week internships were given feedback by local businesses that included "adaptability, initiative, politeness, ability to get along with fellow workers, willingness to take advice and orders, ability to work independently without waiting for suggestions, and easier to learn and advance" (Aikin, 1942, pp. 56-57). This feedback was based on the skills needed to be a successful employee in the 1930s, the rubric could be adjusted to skills found critical for success as an employee in the 2030s and beyond to provide current students participating in an internship an understanding of where growth must occur. This discussion would be considered more beneficial than student performance on standards that will not contribute to their livelihood following their education in grades K through 12. The PEA (1942) observed that the mission of some of the schools participating in the eight-year study changed to state "the work of the secondary school is not completed until each student is satisfactorily established in employment or in college" (p. 68). As Dewey (1902) believed students learned greatest when they lived experiences that could develop connections, Aikin (1942) found, "Study of community often creates a strong desire in young people to do something about conditions which they have discovered" (p. 64). Students were motivated when given the opportunity to learn, outside of the school in a way that fostered connections of previous learning as well experiences from the community they lived in.

The heart of higher-order thinking is addressed in the Eight-Year study when identifying "problem solving develops the habit of reflective thinking" (Aikin, 1942, p. 81). Schools participating in the study shifted their perspective on teaching, "Instead of a lesson to be learned, the work is more often a problem to be solved" (Aikin, 1942, p. 81). This shift is in line with the higher level of popular cognitive frameworks such as extended thinking in Webb's Depth of Knowledge. Feedback from one of the participating studies included "Critical or reflective thinking originates with the sensing of a problem" (Aikin, 1942, p. 82). Here the feedback directly from the participating schools mirrors the reflective thinking Dewey termed in *How We Think.* As the importance for preparing students for the workforce, participating schools found the ability to think critically is an exigency for the success of the country to carry on (Aikin, 1942)

> The success of democracy depends to a large extent on the disposition and ability of citizens to think critically and reflectively about the problems which must of necessity confront them and to improve the quality of their thinking is one of the major goals of education. (p. 82)

In between the publication of *The Cardinal Principles of Secondary Education* and the *Eight-year Study*, Thorndike (1924) published his findings from 1922 and 1923 which consisted of 8,564 students in grades 9, 10, and 11. During these studies, Thorndike (1924) found that students who studied Stenography and typewriting gained more than a student who studied Latin during the same two years. Thorndike (1924) did conclude by describing what a psychologist who was completely separate from the study would find when analyzing the results by writing

“The one causal factor which he would be sure was at work would be the intellect already existent. Those who have the most to begin with gain the most during the year” (p. 13). With a statistical finding as strong as this, educators are charged with the additional task of reducing and ultimately eliminating these gaps to maximize growth for all learners. Thorndike (1924) closed with the finding “After positive correlation of gain with initial ability is allowed for, the balance in favor of any study is certainly not large” (p. 16). This finding leads every educator to ask if the course is not the key to developing the skill set, what is? The answer to this question is found through the different levels of cognitive complexity afforded to students through authentic assessment and the questions being asked of students. Thorndike (1924) concluded his findings by writing:

> “When the good thinkers studied Greek and Latin, these studies seemed to make good thinking. Now that the good thinkers study Physics and Trigonometry, these seem to make good thinking. If the abler pupils should all study Physical Education and Dramatic Art, these subjects would seem to make good thinkers.” (p. 98)

This finding strengthens the claim of Dewey, it is not the courses the students are exposed to that make the thinker, it is the experiences, as well as the level of questioning, opportunity for reflection, and ability to learn how to think and critically think that develops the capacity for students to be able to think.

Runco and Chand's (1995) findings parallel those of Thorndike (1924) when writing about the concept of functional fixedness and creativity. Thorndike’s (1924) finding that the “abler pupils” make “good thinkers” independent of the courses taken can be extrapolated that the students would not only score higher on the general intelligence exam that was used to determine student achievement, but also these students would be considered more creative.

Runco and Chand (1995) defined functional fixedness as “the rigidity or mental set which locks thinking so that an individual cannot see alternatives (p. 247). Functional fixedness “limits thinking” (Runco and Chand, 1995, p. 247) teaching students’ cognitive flexibility will counteract the concerns and limitations of functional fixedness. The lower level thinking tasks that align with Levels 1 and 2 of Webb’s Depth of Knowledge, often found on standardized assessments, are a product of functional fixedness. Questions and tasks that align with recall and remembering are examples of students learning a process, procedure, or set of knowledge from which there is no deviation when regurgitating the information. Runco and Chand (1995) wrote, “if an individual simply retrieves information from memory, it will very probably be unoriginal and thus uncreative (p. 248). Mumford et al. (1993) had the same finding when writing “except under conditions where extant categories are applied in a new situation, the rote application of categories provided by past experience is unlikely to result in novel problem solutions” (p. 128). Sforza et al., (2016) found that “Level 1 and 2 Depth of Knowledge (DOK) complexity accounted for 72% of the high school ELA Common Core State Standards.” (p. 18). Questions on the same state assessments coded as Level 3 or 4, considered cognitively complex thinking, were found for 2% of the questions (Sforza, et al., 2016, p. 19).

Bloom published the *Taxonomy of Educational Objectives: The Classification of Educational Goals. Handbook I: Cognitive Domain* (Bloom, Engelhart, Furst, Hill, & Krathwohl, 1956) in 1956. At the time, Bloom was the Associate Director of the Board of Examinations of the University of Chicago. Bloom created the Taxonomy of Educational Objectives to reduce the labor of preparing for annual comprehensive exams. Bloom’s Taxonomy is a framework for classifying statements of intended instruction or assessment. Of the six categories, all except application were broken down into subcategories. The most common use of Bloom's Taxonomy is to classify objectives and assessment items to show the breadth between the six categories (Krathwohl, 2002). When analyzing the alignment of objectives and assessment items "Almost always, these analyses have heavily emphasized objectives requiring only recognition or recall of information, objectives that fall in the Knowledge category (Krathwohl, 2002, p. 213). This analysis shows the necessity of revising curricula to levels of more complex thinking. Krathwohl continues, "it is objectives that involve the understanding and use of knowledge, those that would be classified in the categories from Comprehension to Synthesis, that are usually considered the more important goals of education" (Krathwohl, 2002, p. 213).

This publication led to educators determining the cognitive complexity of the objective of a lesson or questions on an assessment. Bloom's taxonomy includes six categories of cognitive skills. Since the original publication, the taxonomy has been differentiated between three lower-level cognitive skills, knowledge, comprehension, and application, and three higher-order cognitive skills, analysis, synthesis, and evaluation (Adams, 2015). The three lower-level cognitive skills can easily be replaced with the technological enhancements of the Fourth Industrial Revolution. Through artificial intelligence, common assessments aligned with this level of cognitive complexity are replaced with matching, such as definitions, knowledge recall, and rote memorization. A common and often quickest form of assessment is the knowledge category. The forms of questions that demonstrate mastery of knowledge "can be assessed by straightforward means, for example, multiple-choice or short-answer questions that require the retrieval of recognition of information" (Adams, 2015, p. 1). Adams (2015) differentiates between knowledge and comprehension, through the qualifying characteristic of students being able to paraphrase information, and compare and contrast items, more cognitive complexity than recall is required. Adams (2015) distinguishes the difference between lower and higher levels of cognitive complexity when students are functioning in terms of analysis, synthesis, and evaluation. Noting the difference in the knowledge, comprehension, and application when writing (Adams, 2015) "distinguishing between fact and opinion and identifying the claims upon which an argument is built require analysis, as does breaking down an information need" (p. 1). Table 1 compares the verbs that are used to determine the different levels of thinking in Bloom's Original Taxonomy (1956) and the Revised Taxonomy 2001).

Table 1

Comparison of Descriptors Associated with the Cognitive Process Dimensions of Bloom's original taxonomy (1956) and the Revised Bloom's Taxonomy of Anderson and Krathwohl (2001)

Bloom's Taxonomy (1956)	Revised Bloom Process Dimensions (2005)
Knowledge Define, duplicate, label, list, memorize, name, order, recognize, relate, recall, reproduce, state	*Remember* Retrieve knowledge from long-term memory, recognize, recall, locate, identify
Comprehension Classify, describe, discuss, explain, express, identify, indicate, locate, recognize, report, restate, review, select, translate	*Understand* Construct meaning, clarify, paraphrase, represent, translate, illustrate, provide examples, classify, categorize, summarize, generalize, infer a logical conclusion (such as from examples given), predict, match similar ideas, explain, compare/contrast, construct models (e.g., cause-effect)
Application Apply, choose, demonstrate, dramatize, employ, illustrate, interpret, practice, schedule, sketch, solve, use, write	*Apply* Carry out or use a procedure in a given situation; carry out (apply to a familiar task) or use (apply) to an unfamiliar task
Analysis Analyze, appraise, calculate, categorize, compare, criticize, discriminate, distinguish, examine, experiment, explain	*Analyze* Break into constituent parts, determine how parts relate, differentiate between relevant and irrelevant, distinguish, focus, select, organize, outline, find coherence, deconstruct (e.g., for bias or point of view)
Synthesis Rearrange, assemble, collect, compose, create, design, develop, formulate, manage, organize, plan, propose, set up, write	*Evaluate* Judge based on criteria, check, detect inconsistencies or fallacies, judge, critique
Evaluation Appraise, argue, assess, choose, compare, defend, estimate, explain, judge, predict, rate, core, select, support, value, evaluate	*Create* Combine elements to form a coherent whole, reorganize elements into new patterns/structures, generate, hypothesize, design, plan, construct, produce for a specific purpose

Note. This table contrasts Bloom's Original Taxonomy (Bloom, et. al, 1956) with Bloom's Revised Taxonomy (Anderson and Krathwohl, 2001). From *Cognitive Rigor: Blending the Strengths of Bloom's Taxonomy and Webb's Depth of Knowledge to Enhance Classroom-level Processes* (p. 3) by Hess, et. al, 2009. Publication 2009 by Hess, et al. Reprinted with permission.

Bloom's Revised Taxonomy (Anderson and Krathwohl, 2001) provided two significant changes to the previous taxonomy. The categories of the Revised Taxonomy are: remember, understand, apply, analyze, evaluate, and create. (Anderson & Krathwohl, 2001) The other component added to the Revised Taxonomy is the distinguishing factors of factual, conceptual, procedural, and metacognitive knowledge aligned with tasks, called the Knowledge dimension. Anderson and Krathwohl (2001) refer to the addition of four components, the Cognitive Process Dimension. Factual knowledge represents basic elements needed to understand vocabulary terms and specific critical details and elements. Conceptual knowledge now is one step deeper as it combines relationships of the elements from factual knowledge within the larger context, and how these components function together. Procedural knowledge builds upon conceptual knowledge as it includes criteria for determining when to use appropriate procedures. Finally, the metacognitive knowledge dimension considers a distinction from cognitive psychology, in which students are aware of what they know. (p. 213) Figure 2 explains how Anderson and Krathwohl structured the four different knowledge dimensions of the Revised Taxonomy.

Figure 2

Structure of the Knowledge Dimension of Bloom's Revised Taxonomy

A. Factual Knowledge – The basic elements that students must know to be acquainted with a discipline or solve problems in it.

Aa. Knowledge of terminology

Ab. Knowledge of specific details and elements

B. Conceptual Knowledge – The interrelationships among the basic elements within a larger structure that enable them to function together.

Ba. Knowledge of classifications and categories

Bb. Knowledge of principles and generalizations

Bc. Knowledge of theories, models, and structures

C. Procedural Knowledge – How to do something; methods of inquiry, and criteria for using skills, algorithms, techniques, and methods.

Ca. Knowledge of subject-specific skills and algorithms

Cb. Knowledge of subject-specific techniques and methods

Cc. Knowledge of criteria for determining when to use appropriate procedures

D. Metacognitive Knowledge – Knowledge of cognition in general as well as awareness and knowledge of one's own cognition.

Da. Strategic knowledge

Db. Knowledge about cognitive tasks, including appropriate contextual and conditional knowledge

Dc. Self-knowledge

Note. This figure was produced by Krathwohl in 2002 illustrating the Knowledge Dimension of Bloom's Revised Taxonomy by Anderson and Krathwohl, 2001. From "A Revision of Bloom's Taxonomy: An Overview," by David R. Krathwohl, 2002, *Theory Into Practice*, 41(4), p. 214, Copyright 2002 reprinted with permission of Informa UK Limited, trading as Taylor & Taylor & Francis Group, http://www.tandfonline.com..

The Knowledge dimension of Bloom's Revised Taxonomy aligns with Dewey's belief in thinking as described in *How We Think*. Students operating in the metacognitive knowledge domain align with the form of thinking Dewey would consider reflective thinking. As Dewey (1910) wrote, "reflection involves not simply a sequence of ideas, but a consequence– a consecutive ordering in such a way that each determines the next as its' proper outcome" (p. 2). Similarly, factual knowledge (Anderson and Krathwohl) aligns with belief without an attempt to the state the grounds that support it (Dewey, 1910), which is the opposite of critical thinking, simply accepting fact without analysis or reflection. Dewey (1902) wrote of the importance of metacognitive knowledge when writing, “ To possess all the world of knowledge and lose one’s own self is as awful a fate in education…self-realization is the goal. Learning is active. It involved reaching out of the mind. It involves organic assimilation starting from within. (p. 9)

Webb’s Depth of Knowledge is derived from an analysis of standards and assessments that began in 1997 (Webb, 1999). In 1998, thirteen reviewers analyzed the alignment of assessments and standards in mathematics and science. The results of this analysis provided that alignment between assessments and standards varied. The analysis focused on categorical occurrence, depth-of-knowledge consistency, range-of-knowledge correspondence, and balance of representation. (Webb, 1999, p. 1). The reviewers were given four levels to code assessment activities and standards (Webb, 1999) Level 1, Recall, Level 2, Skill/Concept, Level 3, Strategic Thinking, and Level 4, Extended Thinking. Webb (2007) writes, “standards and assessments can be aligned not only on the category of content covered by each but also based on the complexity of knowledge required by each” (p. 11). Table 2 categorizes four DOK levels as well as provides examples of activities aligned with the corresponding level. The difference in the basis of complexity refers to the four levels found in Webb’s Depth of Knowledge.

Webb (2002) categorized Level 1, as receiving and reciting information or using simple skills or abilities. Lessons are on Level 1 when students memorize and understand words or phrases but do not demonstrate analysis of the text. Dewey (1910) warns of this level of learning “in memorizing this simulated cute and dried copy of the logic of an adult, the child generally is induced to stultify his own subtle and vital logical movement” (p. 69). The recall component of education is what prevents students from the ability to think critically, Webb's Depth of Knowledge (2002), Level 2, extends beyond Level 1, as "Some important concepts are covered, but not in a complex way" (p. 1). Examples of Level 2 thinking are summarizing, interpreting, inferring, and classifying. Main ideas are stressed, and students may apply some of the skills that would be considered Level 1. (Webb, 2002) From lessons aligned with lower levels of thinking, such as what is used in Level 1 or 2 DOK, students will develop "Lack of interest in study, habits of inattention, and procrastination, positive aversion to intellectual application, dependence upon sheer memorizing and mechanical routine with only a modicum of understanding" (Dewey, 1910, pp. 69-70).

Level 3, strategic thinking, and Level 4 extended thinking of Webb's DOK (1997) align with the concept of reflective thinking as described by Dewey (1910). Webb (2007) writes Level 3 thinking requires an explanation of thinking based on reasoning, planning, using evidence, and a higher level of thinking (p. 12). Level 4 thinking requires complex reasoning, planning, developing, and thinking. (p. 12). Dewey (1910) writes:

> All knowledge, all science, thus aims to grasp the meaning of objects and events, and this process always consists in taking them out of their apparent brute isolation as events, and finding them to be parts of some larger whole suggested by them, which, in turn, accounts for, explains, interprets, them. (p. 140)

Therefore, the thinking achieved when working on the DOK level of 3 and 4 is equivalent to what Dewey (1910) writes is "our intellectual life" (p. 140).

Table 2

Webb's Depth-of-Knowledge (DOK) Levels (Webb 1997, 1999)

Level	Description
DOK-1	Recall & Reproduction — Recall a fact, term, principle, or concept; perform a routine procedure.
DOK-2	Basic Application of Skills/Concepts — Use information, conceptual knowledge; select appropriate procedures for a task; perform two or more steps with decision points along the way; solve routine problems; organize or display data; interpret or use simple graphs.
DOK-3	Strategic Thinking — Reason or develop a plan to approach a problem; employ some decision-making and justification; solve abstract, complex, or non-routine problems, complex. (DOK-3 problems often allow more than one possible answer.)
DOK-4	Extended Thinking — Perform investigations or apply concepts and skills to the real world that require time to research, problem solve, and process multiple conditions of the problem or task; perform non-routine manipulations across disciplines, content areas, or multiple sources.

Note. This table describes Webb's Depth of Knowledge (DOK) levels (Webb 1997, 1999) Adapted from Hess, et. al (2009) From *Cognitive Rigor: Blending the Strengths of Bloom's Taxonomy and Webb's Depth of Knowledge to Enhance Classroom-level Processes* (p. 3) by Hess, et. al, 2009. Publication 2009 by Hess, et al. Reprinted with permission.

Hess (2005) developed the Cognitive Rigor Matrix (CRM), after identifying the inconsistencies when using either Bloom's Revised Taxonomy or Webb's Depth of Knowledge. Describing the difference between Bloom's Revised Taxonomy and Webb's Depth of Knowledge, Hess (2013) explains the need for the CRM as:

> Bloom's Taxonomy categorizes the "type of thinking processes" necessary to answer a question. On the other hand, Webb's Depth of Knowledge relates more closely to the depth of content understanding and scope of a learning activity, which manifests in the skills required to complete a task from inception to finale (e.g., planning, researching, drawing conclusions). Each intersection of Bloom-Webb in the matrix provides a focus on differing complexity. (p. 1)

The Cognitive Rigor matrix categorizes learning activities as tools for curriculum development and assessment creation that can have an impact at the classroom level. (Hess, et al., 2009). Cognitive rigor is explained by Hess (2013) as "the complexity of content, the cognitive engagement with that content, and the scope of the planned learning activities" (p. 1). Figures 3 and 4 provide subject-specific examples of Hess's Cognitive Rigor Matrix to determine the cognitive rigor to which students are exposed. Dewey (1902) warns of subject-specific criteria, "Facts are torn away from their original place in experience and rearranged with reference to some general principle. Classification is not a matter of child experience; things do not come to the individual pigeonholed" (p. 7). Hess's Cognitive Rigor Matrix is a matrix that provides the opportunity to identify which cell each question, activity, and assessment aligns with in regard to cognitive rigor, ultimately gauging a balance and structure of cognitive rigor throughout the school day, units, and the entire school year. Hess (2013) wrote, "The Cognitive Rigor Matrix can serve as a constant reminder to educators that students need exposure to novel and complex activities every day" (p.1). Dewey (1902) believes students must be taught how to think and problem solve writing, "The child is expected to "develop" this or that fact or truth out of his own mind. He is told to think things out…without being supplied any of the environing conditions which are requisite to start and guide thought" (p. 16). Providing students with factual knowledge, ultimately leading to extended thinking develops the ability to think critically. Without providing students the time and opportunity in lessons to reason, these skills are not developed, and lessons are taught in a manner of items to be memorized (Dewey, 1902).

Figure 3

Hess's Cognitive Rigor Matrix (CRM) for Reading and Listening Examples

TOOL 1

HESS COGNITIVE RIGOR MATRIX | READING-LISTENING CRM

Integrating Depth-of-Knowledge Levels with Bloom's Cognitive Process Dimensions

Revised Bloom's Taxonomy	DOK Level 1 Recall and Reproduction	DOK Level 2 Skills and Concepts	DOK Level 3 Strategic Thinking or Reasoning	DOK Level 4 Extended Thinking
Remember Retrieve knowledge from long-term memory, recognize, recall, locate, identify	o Recall, recognize, or locate basic facts, terms, details, events, or ideas explicit in texts o Read words orally in connected text with fluency and accuracy	Use these Hess CRM curricular examples with most close reading or listening assignments or assessments in any content area.		
Understand Construct meaning, clarify, paraphrase, represent, translate, illustrate, give examples, classify, categorize, summarize, generalize, infer a logical conclusion, predict, compare–contrast, match like ideas, explain, construct models	o Identify or describe literary elements (characters, setting, sequence, etc.) o Select appropriate words when intended meaning or definition is clearly evident o Describe or explain who, what, where, when, or how o Define or describe facts, details, terms, principles o Write simple sentences	o Specify, explain, show relationships; explain why (e.g., cause-effect) o Give non examples or examples o Summarize results, concepts, ideas o Make basic inferences or logical predictions from data or texts o Identify main ideas or accurate generalizations of texts o Locate information to support explicit-implicit central ideas	o Explain, generalize, or connect ideas using supporting evidence (quote, example, text reference) o Identify or make inferences about explicit or implicit themes o Describe how word choice, point of view, or bias may affect the readers' interpretation of a text o Write multi paragraph composition for specific purpose, focus, voice, tone, and audience	o Explain how concepts or ideas specifically relate to other content domains (e.g., social, political, historical) or concepts o Develop generalizations of the results obtained or strategies used and apply them to new problem-based situations
Apply Carry out or use a procedure in a given situation; carry out (apply to a familiar task), or use (apply) to an unfamiliar task	o Use language structure (pre-, or suffix) or word relationships (synonym or antonym) to determine meaning of words o Apply rules or resources to edit spelling, grammar, punctuation, conventions, word use o Apply basic formats for documenting sources	o Use context to identify the meaning of words or phrases o Obtain and interpret information using text features o Develop a text that may be limited to one paragraph o Apply simple organizational structures (paragraph, sentence types) in writing	o Apply a concept in a new context o Revise final draft for meaning or progression of ideas o Apply internal consistency of text organization and structure to composing a full composition o Apply word choice, point of view, style to impact readers' or viewers' interpretation of a text	o Illustrate how multiple themes (historical, geographic, social, artistic, literary) may be interrelated o Select or devise an approach among many alternatives to research a novel problem
Analyze Break into constituent parts, determine how parts relate, differentiate between relevant-irrelevant, distinguish, focus, select, organize, outline, find coherence, deconstruct (e.g., for bias or point of view)	o Identify whether specific information is contained in graphic representations (e.g., map, chart, table, graph, T-chart, diagram) or text features (e.g., headings, subheadings, captions) o Decide which text structure is appropriate to audience and purpose	o Categorize or compare literary elements, terms, facts or details, events o Identify use of literary devices o Analyze format, organization, and internal text structure (signal words, transitions, semantic cues) of different texts o Distinguish: relevant-irrelevant information; fact or opinion o Identify characteristic text features; distinguish between texts, genres	o Analyze information within data sets or texts o Analyze interrelationships among concepts, issues, problems o Analyze or interpret author's craft (literary devices, viewpoint, or potential bias) to create or critique a text o Use reasoning, planning, and evidence to support inferences	o Analyze multiple sources of evidence, or multiple works by the same author, or across genres, time periods, themes o Analyze complex or abstract themes, perspectives, concepts o Gather, analyze, and organize multiple information sources o Analyze discourse styles
Evaluate Make judgments based on criteria, check, detect inconsistencies or fallacies, judge, critique	"UG"—unsubstantiated generalizations = stating an opinion without providing any support for it!		o Cite evidence and develop a logical argument for conjectures o Describe, compare, and contrast solution methods o Verify reasonableness of results o Justify or critique conclusions drawn	o Evaluate relevancy, accuracy, and completeness of information from multiple sources o Apply understanding in a novel way, provide argument or justification for the application
Create Reorganize elements into new patterns or structures, generate, hypothesize, design, plan, produce	o Brainstorm ideas, concepts, problems, or perspectives related to a topic, principle, or concept	o Generate conjectures or hypotheses based on observations or prior knowledge and experience	o Synthesize information within one source or text o Develop a complex model for a given situation o Develop an alternative solution	o Synthesize information across multiple sources or texts o Articulate a new voice, alternate theme, new knowledge or perspective

Note. Hess's Cognitive Rigor Matrix for Reading and Listening. From Cognitive Rigor & DOK Focus Area by K.K. Hess, 2009, Karin-Hess.com (https://www.karin-hess.com/_files/ugd/5e86bd_1c0a13cbb2bc4f1185a558dbfeb27ffd.pdf). Copyright 2009 by Karin Hess. Reprinted with permission.

Figure 4

Hess's Cognitive Rigor Matrix (CRM) for Math and Science Examples

TOOL 2

HESS COGNITIVE RIGOR MATRIX | MATH-SCIENCE CRM

Integrating Depth-of-Knowledge Levels with Bloom's Cognitive Process Dimensions

Revised Bloom's Taxonomy	DOK Level 1 Recall and Reproduction	DOK Level 2 Skills and Concepts	DOK Level 3 Strategic Thinking or Reasoning	DOK Level 4 Extended Thinking
Remember Retrieve knowledge from long-term memory, recognize, recall, locate, identify	o Recall, observe, and recognize facts, principles, properties o Recall/ identify conversions among representations or numbers (e.g., customary and metric measures)	Use these Hess CRM curricular examples with most mathematics or science assignments or assessments.		
Understand Construct meaning, clarify, paraphrase, represent, translate, illustrate, give examples, classify, categorize, summarize, generalize, infer a logical conclusion, predict, compare–contrast, match like ideas, explain, construct models	o Evaluate an expression o Locate points on a grid or number on number line o Solve a one-step problem o Represent math relationships in words, pictures, or symbols o Read, write, compare decimals in scientific notation	o Specify and explain relationships (e.g., non examples or examples; cause-effect) o Make and record observations o Explain steps followed o Summarize results or concepts o Make basic inferences or logical predictions from data or observations o Use models or diagrams to represent or explain mathematical concepts o Make and explain estimates	o Use concepts to solve non routine problems o Explain, generalize, or connect ideas using supporting evidence o Make and justify conjectures o Explain thinking or reasoning when more than one solution or approach is possible o Explain phenomena in terms of concepts	o Relate mathematical or scientific concepts to other content areas, other domains, or other concepts o Develop generalizations of the results obtained and the strategies used (from investigation or readings) and apply them to new problem situations
Apply Carry out or use a procedure in a given situation; carry out (apply to a familiar task), or use (apply) to an unfamiliar task	o Follow simple procedures (recipe-type directions) o Calculate, measure, apply a rule (e.g., rounding) o Apply algorithm or formula (e.g., area, perimeter) o Solve linear equations o Make conversions among representations or numbers, or within and between customary and metric measures	o Select a procedure according to criteria and perform it o Solve routine problem applying multiple concepts or decision points o Retrieve information from a table, graph, or figure and use it solve a problem requiring multiple steps o Translate between tables, graphs, words, and symbolic notations (e.g., graph data from a table) o Construct models given criteria	o Design investigation for a specific purpose or research question o Conduct a designed investigation o Use concepts to solve non routine problems o Use and show reasoning, planning, and evidence o Translate between problem and symbolic notation when not a direct translation	o Select or devise approach among many alternatives to solve a problem o Conduct a project that specifies a problem, identifies solution paths, solves the problem, and reports results
Analyze Break into constituent parts, determine how parts relate, differentiate between relevant-irrelevant, distinguish, focus, select, organize, outline, find coherence, deconstruct	o Retrieve information from a table or graph to answer a question o Identify whether specific information is contained in graphic representations (e.g., table, graph, T-chart, diagram) o Identify a pattern or trend	o Categorize, classify materials, data, figures based on characteristics o Organize or order data o Compare–contrast figures or data o Select appropriate graph and organize and display data o Interpret data from a simple graph o Extend a pattern	o Compare information within or across data sets or texts o Analyze and draw conclusions from data, citing evidence o Generalize a pattern o Interpret data from complex graph o Analyze similarities–differences between procedures or solutions	o Analyze multiple sources of evidence o Analyze complex or abstract themes o Gather, analyze, and evaluate information
Evaluate Make judgments based on criteria, check, detect inconsistencies or fallacies, judge, critique	"UG"—unsubstantiated generalizations = stating an opinion without providing any support for it!		o Cite evidence and develop a logical argument for concepts or solutions o Describe, compare, and contrast solution methods o Verify reasonableness of results	o Gather, analyze, and evaluate information to draw conclusions o Apply understanding in a novel way, provide argument or justification for the application
Create Reorganize elements into new patterns or structures, generate, hypothesize, design, plan, produce	o Brainstorm ideas, concepts, or perspectives related to a topic	o Generate conjectures or hypotheses based on observations or prior knowledge and experience	o Synthesize information within one data set, source, or text o Formulate an original problem given a situation o Develop a scientific/mathematical model for a complex situation	o Synthesize information across multiple sources or texts o Design a mathematical model to inform and solve a practical or abstract situation

Note. Hess's Cognitive Rigor Matrix for Reading and Listening. From Cognitive Rigor & DOK Focus Area by K.K. Hess, 2009, Karin-Hess.com (https://www.karin-hess.com/_files/ugd/5e86bd_db128ad5d1a44f549bc5cc12c15b9799.pdf). Copyright 2009 by Karin Hess. Reprinted with permission.

The highest level of thinking through the revised cognitive framework of Bloom is creation, this level of thinking corresponds with the action words synonymous with Webb's Extended Thinking category, such as, arrange, collect, construct, design, develop, formulate, organize, and create.. Bloom's categories of application analysis, evaluate, create, and Webb's Levels 3 and 4 parallel the type of thinking that Dewey claims is "the business" of school, ultimately teaching individuals to think critically and creatively (Dewey, 1910). The theoretical framework utilized throughout the study will be to identify if the coding of questions in Hess' Cognitive Rigor Matrix aligns with the experimentalism philosophy of education identified by Dewey, Thorndike, Aikin, and the Commission on the Reorganization of Secondary Education.

Dating back to the end of the 19th century, the progressive philosophy of education has demonstrated the greatest possibility of retention of information in combination with the ability of students to transfer knowledge when in different contexts. Dewey has delineated examples of instructional practices which align with the ability to develop critical thinking skills. These critical thinking skills rank among employees' identified needs for 2030 and beyond based on surveys from leading international organizations. Only through developing these skills and building a foundation in the classrooms can teachers and administrators fulfill the expectation of preparing our students to be successful beyond their time in the schoolhouse. Research indicates the onslaught of standardized testing often contradicts the development of higher-order, cognitively complex thinking. With standardized testing being scored by artificial intelligence and simple scantron, answers are graded at rapid speeds, which do not account for critical thinking. To camouflage this deficiency, standardized tests have been written in difficult to understand syntax in an attempt to norm student achievement and overall performance scores. The use of standardized tests contradicts findings from seminal works such as T*he Cardinal Principles of Secondary Education, The Eight-Year Study,* and current research by not incorporating the ability to assess multi-step assessments in which students are required to demonstrate the ability to conduct cognitively complex tasks. Additionally, standardized tests do not address the insufficiencies that have been identified by educational researchers such as Thorndike.

The need for consistency in the ability has been identified through several cognitive frameworks. This study will analyze if the CCRA+ provided by the CAE is a standardized assessment that will assess the ability of students to think critically, problem solve and effectively write as claimed through the utilization of Hess' Cognitive Rigor Matrix, which overlaps Bloom's Revised Taxonomy and Webb's DOK.

Chapter III will include an analysis of the methodology for this study which includes an introduction to the study, the research questions guiding the study, and a description of the design and purpose of the study. The qualifications and coding process will be chronicled, as well as how reliability and validity were met throughout the study. Chapter III will also include an explanation of how the mixed-methods study was conducted as well as how the level of cognitive complexity on the questions of the CCRA+ align with Hess' Cognitive Rigor Matrix.

Chapter III

Methodology

Introduction

This study aims to categorize the cognitive complexity and frequency of each level of complexity of the questions found on the CCRA+ assessment based on Hess' Cognitive Rigor Matrix categories. The CCRA+ provides one assessment for middle school students in grades 6-8 and one assessment for high school students in grades 9-12. This study will analyze the cognitive complexity of retired, publicly available questions from one middle school assessment and from two high school student guides and a presentation on the high school CCRA+ assessment. According to the test developers, the publicly available questions are similar in complexity to the questions on the currently used assessments.

The test developer, the Council for Aid to Education (CAE), claims the CCRA+ assessment can provide feedback to determine if students are prepared to be successful in either College or a Career after completing the twelfth grade. According to CAE, critical thinking, problem-solving, and effective written communication skills are assessed authentically to gauge students' level or preparedness. This study used Hess' Cognitive Rigor Matrix as an analytical framework to determine the complexity of the questions assessing these skills. This chapter details the methodology of how this study was conducted and how the study findings were determined.

Research Questions

The study was guided by the overarching research question: What type of thinking is required by the College and Career Readiness Assessment (CCRA+) for middle and high school students?

Two sub-questions guided the study:

a.) What is the frequency and percentage of questions categorized as higher-level within each cell of Hess' Cognitive Rigor Matrix on the CCRA+?

b.) In what ways are the skills identified as essential for success in the workplace by global organizations assessed in the CCRA+?

The CCRA+ comprises two sections, a performance assessment task that includes document analysis from real documents used to develop an argument followed by a constructed response and selected-response questions that have sub-foci on data literacy, critical reading, and critiquing an argument.

Research Design and Methodology

A convergent, parallel mixed-methods study with qualitative content analysis and quantitative descriptive statistics was conducted to describe how the language of the questions on the CCRA+ test compares with the language associated with higher-order thinking found in the research literature and (b) to describe how the skills identified as essential for success in the workplace by global organizations are assessed in the CCRA+.

This study included deductive category application to determine the cognitive complexity of 20 questions found on a CCRA+ middle school assessment, student guides, and information session. The deductive application uses previously developed frameworks in connection with the text that will be analyzed based on a predetermined coding scheme (Mayring, 2000). For this study, the text analyzed will be retired, publicly available CCRA+ assessment prompts analyzed through Hess' Cognitive Rigor Matrix framework. The categories of cognitive complexity were pre-existing. Hess' Cognitive Rigor Matrix is one appropriate framework for this study as it analyzes the cognitive rigor of assessments on each question based on two previously found valid and reliable frameworks, Webb's DOK and Bloom's Revised Taxonomy (Hess et al., 2009).

Reliability

Reliability was reached through expert coder training following the Webb Alignment Tool training manual (Webb, et al., 2005). Hess' Cognitive Rigor Matrix is reliable as the two superimposed documents to create the matrix. Bloom's Revised Taxonomy and Webb's DOK are widely used and considered credible. The study used three expert coders who analyzed the cognitive complexity of the assessment questions. Each coder read and independently categorized each question per Hess' Cognitive Rigor Matrix and the Webb Alignment Tool. Next, coders compared their categorization for inter-rater reliability, as Merriam (2009) suggested.

The American Psychological Association (APA) defines inter-rater reliability and explains its importance as "a measure of consistency used to evaluate the extent to which different judges agree in their assessment decisions. Inter-rater reliability is essential when making decisions in research or clinical settings" (APA, 2010, para. 1). The work of Miles, et al., (2014) was utilized to increase reliability based on the double-rater read behind consensus model.

Webb (1997) developed an alignment tool between standards, standardized assessments, and curriculum. Webb's Depth of Knowledge (DOK). Through an analysis to align expectations and assessment in the same school, across districts, and states, Webb, aligned tests with a "curriculum based on compatible and sound principles of cognitive development" (Webb, 1997, p. 12). Webb's DOK focuses on determining the cognitive complexity of questions, not the difficulty of the questions on an assessment. Webb's DOK is widely used throughout education to critique the level of complexity that is being addressed in assessments and curricula. Progressing from DOK Level 1 to 4, each level contains a deeper level of thinking to answer the question or complete the task successfully. Webb (1997) includes all forms of knowledge, such as procedural and declarative in thought throughout the different levels.

Demonstrating knowledge according to Webb's four levels of DOK assesses learning and understanding similar to a ladder. To demonstrate the deepest level of thinking, Level 4, a learner must transfer previously acquired knowledge from less cognitively demanding levels. Webb specified differences to observe and expect among the four content areas of English/language arts which were further separated to reading and writing, mathematics, science, and social studies. The four levels of Webb's DOK are described below as written in Web Alignment Tool (WAT) Training Manual (2005, pp. 73-74):

Level 1 (Reading) - Requires students to receive or recite facts or to use simple skills or abilities

Level 2 (Reading) - Includes the engagement of some mental processing beyond recalling or reproducing a response; it requires both comprehension and subsequent processing of text or portions of text.

Level 3 (Reading) - Deep knowledge becomes a greater focus. Students are encouraged to go beyond the text; however, they are still required to show understanding of the ideas in the text. Students may be encouraged to explain, generalize, or connect ideas.

Level 4 (Reading) - Higher-order thinking is central, and knowledge is deep. The standard or assessment item at this level will probably be an extended activity with extended time provided for completing it. The extended time period is not a distinguishing factor if the required work is only repetitive and does not require the application of significant conceptual understanding of higher-order thinking.

Bloom's original taxonomy (1956) was revised in 2001 by Anderson and Krathwohl. The most significant adaptation to the original taxonomy is the revised taxonomy "intersects and acts upon different types and levels of knowledge – factual, conceptual, procedural, and metacognitive" (Wilson, 2016, p. 2). Below is a description of the six levels of the Revised Taxonomy as written by Krathwohl (2002)

Remember- Retrieving relevant knowledge from long-term memory

Understand- Determining the meaning of instructional messages, including oral, written, and graphing communication

Apply- Carrying out or using a procedure in a given situation

Analyze- Breaking material into its constituent parts and detecting how the parts relate to one another and to an overall structure or purpose

Evaluate- Making judgments based on criteria and standards

Create- Putting elements together to form a novel, coherent whole or make an original product.

Data Collection

Qualitative Data

The qualitative data for this study were found through publicly available performance tasks and selected-response tasks from previous CCRA+ assessments from the middle school and high school levels. In total, three performance tasks and seventeen selected-response questions were gathered for analysis. For the process of best understanding the assessment, including its reliability and validity, the researcher had phone calls with representatives from CAE.

Due to the limited number of retired, publicly accessible questions, each provided question was analyzed. The questions that were analyzed are written below:

- CCRA+ Middle School Performance Task: 1 Question
- CCRA+ Middle School Selected Response: 13 Questions
- CCRA+ High School Performance Tasks: 2 Questions
- CCRA+ High School Selected Response: 4 Questions

The CCRA+ assessment is composed of selected-response questions that can only be answered through the provided documents in the assessment. The selected-response questions are separated into three categories, data literacy, critical reading, and critiquing an argument. Students are given 30 minutes to complete the selected-response question section of the test. Students are given 60 minutes to complete one performance task. The performance task is a written, constructed response that requires students to cite information from multiple documents to create and synthesize a logical response to a real-world problem. There are several correct responses to the performance task based on the information the students select from the provided sources. Both the selected-response and constructed response questions are answered based on documents embedded in the assessment, such as office memorandum, letters, press releases, maps, data tables, technical reports, e-mails, and maps (CAE). The sampling of the 17 selected-response questions and three performance tasks will be coded, which provides a sample size of 20 questions and one hundred percent (100%) of the total number of questions available for analysis.

Quantitative Data

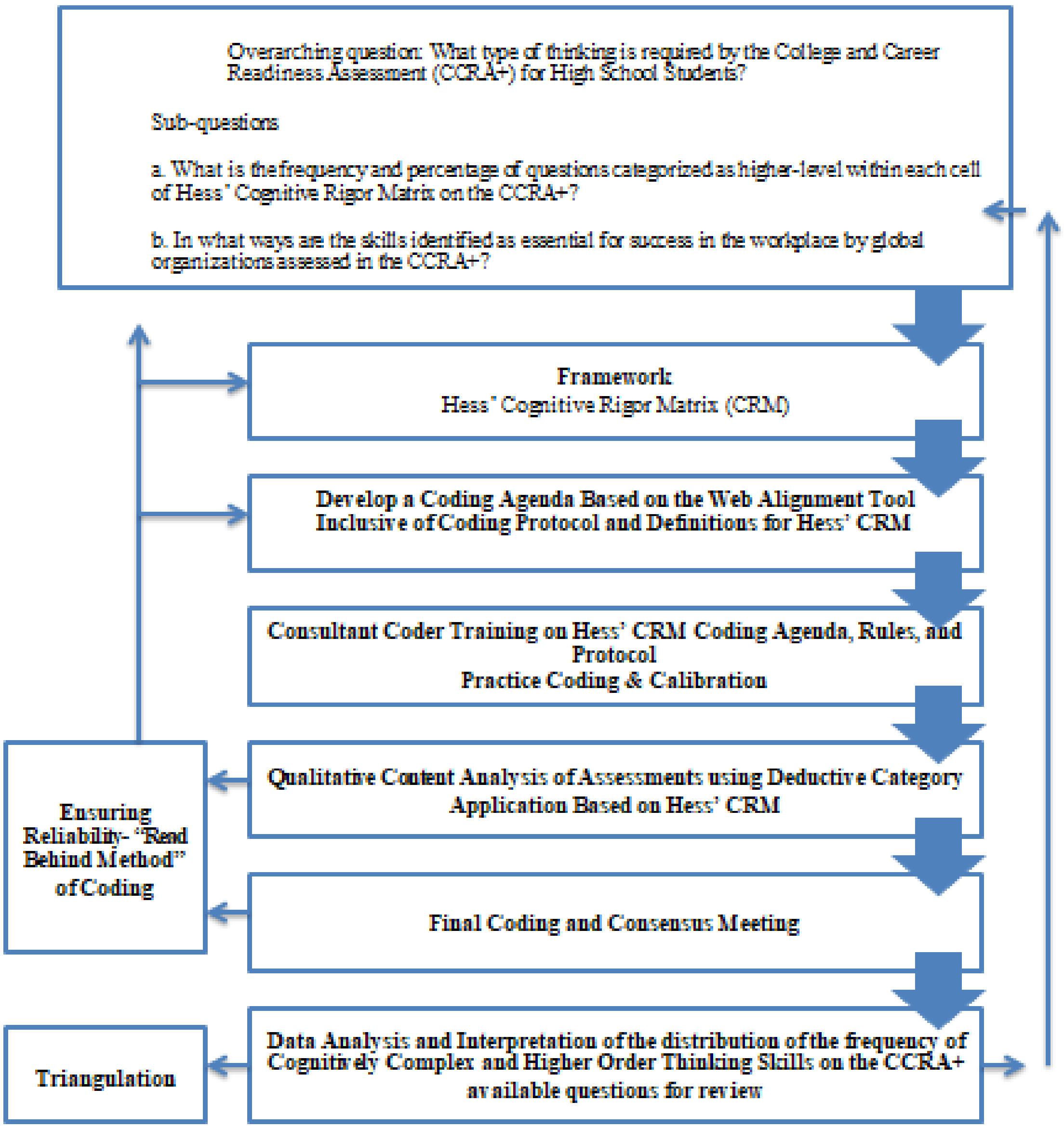
The quantitative data were gathered concurrently with the analysis of the cognitive complexity of the qualitative data. While coding the questions per Hess' Cognitive Rigor Matrix, a frequency tally of each aligning cell was collected. The frequency was calculated by initially gathering the total number as a raw number. The raw number was then converted to a percentage. Questions coded in the category of Webb's DOK at the third or fourth level were considered of high cognitive complexity, and questions coded in Webb's DOK at level one or two were considered low cognitive complexity, as agreed upon by the coders and research literature.

Methods

The researcher conducted a convergent, parallel mixed-methods study with qualitative content analysis methods and quantitative descriptive statistics. Hess' Cognitive Rigor Matrix categorized 20 questions from retired, publicly available CCRA+ assessment prompts to use deductive category application. Qualitative methods were utilized to determine the cognitive complexity of the questions asked on the CCRA+ assessments per Hess' Cognitive Rigor Matrix. Quantitative methods were utilized in the study to identify the frequency of the questions that were coded as either higher level or lower level cognitive complexity. The figure below illustrates Mayring's Step Model (2000), which describes this study's deductive application process and triangulation of findings.

Figure 5

Step Model for Deductive Category Application, adapted from Mayring (2000)



Note. The step model above illustrates the process of deductive category application for this study. Adapted from "Qualitative Content Analysis" by P Mayring, 2000, *Forum Qualitative Social Research, 1*(2), p. 4 (https://www.qualitative-research.net/index.php/fqs/article/view/1089/2386) Copyright 2000 by FQS

Qualitative Methods

Hess' Cognitive Rigor Matrix is a tool to analyze the complexity of questions on the CCRA+. Bloom's Revised Taxonomy, classifies the type of thinking required to answer a question, whereas Webb's DOK determines the depth of content understanding and skills required to answer a question (Hess, 2013) successfully. Hess' Cognitive Rigor Matrix will allow for a reference of type and depth of thinking required of students to answer each question.

Quantitative Methods

A frequency chart was completed as each question was read and categorized per Hess' Cognitive Rigor Matrix. The data from the frequency chart were then converted from raw numbers to percentages based on what has been determined by research literature and the coders as higher cognitive complexity Webb's Level 3 and Webb's Level 4 DOK, along with the aligned Bloom's Revised Taxonomy levels. Lower-level cognitive complexity was questions found to align with cells that had questions from Webb's Level 1 and Webb's Level 2 DOK along with the aligned Bloom's Revised Taxonomy levels. The total number of questions considered higher-level cognitive complexity and the total number of questions considered lower level cognitive complexity was the numerator, and the total number of sample questions (20) was the denominator for calculating both levels of cognitive complexity.

Consultant Coder

Three trained coders were asked to code the questions analyzed on the CCRA+ to reach a consensus on the level of cognitive complexity as determined by Hess' Cognitive Rigor Matrix. The coders consisted of a middle school teacher, an elementary school teacher, and the researcher. Both coders have earned a Doctor of Education degree in Education Leadership, Management, and Policy. Both coders have expertise in the fields of cognitive complexity and in online and standardized assessments, which proved essential in the coding process for this study.

Coding Scheme

A noteworthy difference between Bloom’s Revised Taxonomy and Webb’s DOK is that tasks and questions are not necessarily cognitively complex at any level of Bloom’s Revised Taxonomy. Hess et al, (2009) identified, “The resulting combination of Bloom’s Taxonomy and depth of knowledge – cognitive rigor – forms a comprehensive structure for defining rigor, thus posing a wide range of uses at all levels of curriculum development and instruction. (p. 1) Hess’ Cognitive Rigor Matrix superimposes Bloom’s Revised Taxonomy and Webb’s Depth of Knowledge, ultimately providing educators the ability to determine a specific cell in which the cognitive complexity of each question or task can be determined. Hess, et al., (2009) found, “The resulting cognitive rigor matrix… vividly connects, yet clearly distinguishes, the two schemata, allowing educators to examine the rigor associated with tasks that might seem at first glance comparable in complexity” (p. 5). The use of Hess’ Cognitive Rigor Matrix is the most precise framework for this study.

Hess’ Cognitive Rigor Matrix is created with Webb’s DOK listed horizontally and Bloom’s Revised Taxonomy written vertically. There are five vertical columns along the horizontal axis of Hess’ Cognitive Rigor Matrix. The furthest left column includes the levels of Bloom’s Revised taxonomy followed by the four levels of Webb’s DOK. Webb’s DOK is described as Level 1 Recall and Reproduce, Level 2 Skills and Concepts, Level 3 Strategic Thinking/Reasoning, Level 4 Extended Thinking. Horizontally, the first row states Remember: Retrieve knowledge from long-term memory, recognize, recall, locate, identify. The second row begins with, Understand: construct meaning, clarify, paraphrase, represent, translate, illustrate, give examples, classify, categorize, summarize, generalize, infer a logical conclusion, predict, compare/contrast, match like ideas, explain, construct models. The third-row defines Apply as: Carry out or use a procedure in a given situation; carry out (apply to a familiar task) or use (apply) to an unfamiliar task. Followed by the fourth-row which defines Analyze: Break into constituent parts, determine how parts relate, differentiate between relevant-irrelevant, distinguish, focus, select, organize, outline, first coherence, deconstruct (e.g. for bias or point of view). Next, Evaluate is described as: Make judgments based on criteria, check, detect inconsistencies or fallacies, judge, and critique. Finally, Create is defined as reorganizing elements into new patterns/structures, generating, hypothesizing, designing, planning, constructing, and producing (p. 8). The description that would correlate with a question or task aligning with a particular cell of Hess' Cognitive Rigor Matrix is described below as written by Hess, et al (2009) when listing the complexity of questions in parentheses, Webb's DOK is listed first followed by Bloom's Revised Taxonomy, which is found in Appendix A :

- Webb's Level 1, Bloom's Level 1 (A,1): Recall, recognize, locate basic facts, ideas, principles Recall or identify conversions: between units of measure Identify facts/details in texts
- Webb's Level 1, Bloom's Level 2 (A,2): Compose/decompose numbers Evaluate an expression Locate points on a grid Symbolize math relationships Write simple sentences Describe/explain how or why
- Webb's Level 1, Bloom's Level 3 (A,3): Follow simple/routine procedures Solve a one-step problem Calculate, measure, apply a rule Apply an algorithm or formula Represent in words or diagrams a concept or relationship Apply rules or use resources to edit spelling and grammar

- Webb’s Level 1 Bloom’s Level 4 (A,4): Retrieve information from a table or graph to answer a question Identify or locate specific information contained in maps, charts, tables, graphs, or diagrams
- Webb’s Level 1, Bloom’s Level 5 (Excluded from coding chart): Examples are not included in Hess’ Cognitive Rigor Matrix
- Webb’s Level 1, Bloom’s Level 6 (A, 6): Brainstorm ideas, concepts, or perspectives related to a topic or concept
- Webb’s Level 2, Bloom’s Level 1 (Excluded from coding chart): Examples are not included in Hess’ Cognitive Rigor Matrix
- Webb’s Level 2, Bloom’s Level 2 (B,2): Specify and explain relationships Give non-examples/examples Make and record observations Summarize results, concepts, ideas Infer or predict from data or texts Identify main ideas
- Webb’s Level 2, Bloom’s Level 3 (B,3): Select a procedure according to task needed and perform it Solve routine problem applying multiple concepts or decision points Retrieve information from a graph and use it solve a multi-step problem Use models to represent concepts Write paragraph using appropriate organization, text structure
- Webb’s Level 2, Bloom’s Level 4 (B,4): Categorize, classify materials Compare/ contrast figures or data Select appropriate display data Extend a pattern Identify use of literary devices Identify text structure of paragraph
- Webb’s Level 2, Bloom’s Level 5 (Excluded from coding chart): Examples are not included in Hess’ Matrix

- Webb's Level 2, Bloom's Level 6 (B,6): Generate conjectures or hypotheses based on observations or prior knowledge
- Webb's Level 3, Bloom's Level 1 (Excluded from coding chart): Examples are not included in Hess' Cognitive Rigor Matrix
- Webb's Level 3, Bloom's Level 2 (C,2): Explain, generalize, or connect ideas using supporting evidence Explain phenomena in terms of concepts Write full composition to meet specific purpose Identify themes
- Webb's Level 3, Bloom's Level 3 (C,3): Use concepts to solve non-routine problems Design investigation for a specific purpose or research question Conduct a designed investigation Use reasoning, planning, and evidence Revise final draft for meaning or progression of ideas
- Webb's Level 3, Bloom's Level 4 (C,4): Compare information within or across data sets or texts Analyze and draw conclusions Generalize a pattern Organize/interpret data Analyze author's craft or viewpoint
- Webb's Level 3, Bloom's Level 5 (C,5): Cite evidence and develop a logical argument for concepts Describe, compare, and contrast solution methods Verify reasonableness of results Justify conclusions made
- Webb's Level 3, Bloom's Level 6 (C,6): Synthesize information within one source or text Formulate an original problem Develop a complex model for a given situation
- Webb's Level 4, Bloom's Level 1 (Excluded from coding chart): Examples are not included in Hess' Cognitive Rigor Matrix

- Webb’s Level 4, Bloom’s Level 2 (D,2): Explain how concepts or ideas specifically relate to other content domains or concepts Develop generalizations of the results obtained or strategies used and apply them to new problem situations
- Webb’s Level 4, Bloom’s Level 3 (D,3): Select or devise an approach among many alternatives to solve a novel problem Conduct a project that specifies a problem, identifies solution paths, solves the problem, and reports results Illustrate how multiple themes (historical, geographic, social) may be interrelated
- Webb’s Level 4, Bloom’s Level 4 (D,4): Analyze multiple sources of evidence or multiple works by the same author, or across genres Analyze complex/abstract themes Gather, analyze, and organize information Analyze discourse styles
- Webb’s Level 4, Bloom’s Level 5 (D,5): Gather, analyze, and evaluate relevance and accuracy Draw and justify conclusions Apply understanding in a novel way, provide argument or justification for the application
- Webb’s Level 4, Bloom’s Level 6 (D,6): Synthesize information across multiple sources or texts Design a model to inform and solve a real-world, complex, or abstract situations

Coding Process, Data Analysis, and Reliability

The CCRA+ has a selected response and constructed response performance task component. The three sections of the selected response component are critical reading, data literacy, and critiquing an argument. The English Language Arts version of Hess’ Cognitive Rigor Matrix was utilized for coding. English Language Arts was selected because students must comprehend the questions and documents provided to logically create a constructed response performance task.

As explained by Miles et al. (2014) a double-rater read behind consensus model was used to increase reliability. The researcher was trained on how to code test questions per Webb's Alignment Tool Training Manual (2005) and Hess' Cognitive Rigor Matrix (2009) in a two hour training session conducted by Dr. Christopher Tienken on June 15th, 2022. During the initial training session, the practice of the double-rater read behind consensus model was also modeled and practiced. Dr. Tienken previously trained the additional two coders in the same method. Each analyst used Mayring's Step Model (2000) to obtain data triangulation and consistently guide the data analysis process.

The expert coders and the researcher met on December 1st, 2022 via Zoom (zoom.us) for the calibration session to code questions from Webb's Alignment Tool Training Manual (2005) and three questions from the CCRA+ assessments. During this session, each question was individually analyzed. The coders first analyzed ten questions from Webb's Alignment Tool Training Manual (2005), eight of the ten questions were direct consensus, two of the ten questions reached consensus indirectly following a discussion. Consensus was reached on all ten or 100% of the calibration questions. Following the calibration questions, a selected-response question and two performance tasks from high school assessments were analyzed. Of these questions, direct consensus was reached on the selected-response question and one performance task, indirect consensus was reached on the second performance task. Consensus was reached on all three or 100% of the analyzed questions. Following initial meeting on December 1st, an independent reading, analysis, and coding of each question per set was conducted by the coders. The coders reconvened on December 6th, 2022 via Zoom (zoom.us). During this coding session, the coders tallied their data of the remaining seventeen questions in a table which were held on a Google Sheet. Discussion ensued as necessary if there was a discrepancy in the coding of a cell according to Hess' Cognitive Rigor Matrix. If there was a discrepancy, the question was highlighted to ensure documentation of the discrepancy. The discussion that followed a discrepancy in coding the particular level of cognitive complexity included a rationale from each coder for the assignment of the cognitive complexity aligned with the corresponding cell of the Cognitive Rigor Matrix based on the language and thinking required to successfully answer the question. The use of discussion as required when a discrepancy was identified employed the double-rater read behind consensus model, ultimately increasing inter-rater reliability. The frequency of exact agreement as well as the need for the discussion of discrepancies was calculated and recorded. The data table used to code the cognitive complexity of the analyzed questions was developed by Solis-Stovall (2020). The coding data table for the constructed response performance task questions is below. The horizontal row lists the coordinating cell for Hess' Cognitive Rigor Matrix, for example, A, 1 represents level one of Bloom's Revised Taxonomy, remembering, and level 1 of Webb's Depth of Knowledge, recall. The vertical column indicates the specific question from the assessment, HS-A represents the first constructed response performance task from the High School Assessment, HS-B represents the second constructed response performance task from the High School Assessment, MS represents the only middle school constructed response performance task, results of Figure 6 can be found in Appendix B.

Figure 6

Coding Table for Constructed Response Performance Task Questions. Adapted from Solis-Stovall (2020)

	A,1	A,2	A,3	A,4	A,6	B,2	B,3	B,4	B,6	C,2	C,3	C,4	C,5	C,6	D,2	D,3	D,4	D,5	D,6
HSA																			
HSB																			
MS																			

Note. This table depicts the frequency of Constructed Response questions per cell of Hess' Cognitive Rigor Matrix Adapted from An Analysis of the Higher Order Thinking Requirements of PARCC Practice Assessments in Grades 3 and 4 by L. Solis-Stovall, 2020 (https://scholarship.shu.edu/cgi/viewcontent.cgi?article=3880&context=dissertations).

The same table was replicated for the selected-response questions for both the middle school and high school assessments and can be found in Appendix C. The protocol for the deliberation following coding the three sets of questions from the CCRA+ has been adapted from Vanderhook (2020) and Lamberti (2020) is listed below:

- Coders independently reviewed questions from four sources in one set
- Coders reconvened to review the score of each question in the particular set that was reviewed.
- If coders agreed on the corresponding cognitive rigor, the coders would proceed to the next question
- If there was a disparity in the cognitive rigor based on the language of a question by the coders, a discussion followed.

- If consensus was not met, the questions were coded for the higher level of cognitive rigor per the initial coding. If there was a difference in the coding of Bloom's Taxonomy or Webb's DOK, Webb's DOK was considered more cognitively rigorous.
- Frequency of cognitive complexity was then calculated for that set of questions
- Following the coding of all three sections frequency was determined for each section individually and all three sections were analyzed, in total.

This protocol has previously been modeled in similar studies of cognitive complexity to provide consistency in coding and deliberation as required. (Sato, Lagunoff, & Worth, 2011:Miles, Huberman, & Saldana. 2014; as cited in Sforza, Tienken, & Kim, 2016)

Following the analysis of each set of questions, the coders assessed their responses to ensure reliability. The coders completed 20 questions with 80% direct agreement and 100% consensus by the end of the session. The double-rater read behind method was instilled in any instance that did not lead to agreement in coding the cognitive complexity of a question. In set one, a high school selected-response question and performance task the coders agreed on 50% of the questions, moving the performance task on Hess' Cognitive Rigor Matrix from cell D, 6 to cell D, 5. The coders agreed on 77% of the middle school selected-response questions. Moving only two questions on Hess' Cognitive Rigor Matrix from cell C, 4 to cell B, 2 and moving one question from cell C, 3 to B, 2 following discussion. For both the individual performance task question from a student guide and three selected-response questions from a slideshow presentation, the coders agreed on 100% of the questions. When agreement was not met independently, the double-rater read behind method was utilized to discuss the process of analysis and come to a consensus as appropriate. After the double-rater read behind method was implemented, 100% consensus was met for all questions.

After consensus was met, the coders calculated the frequency and distribution of the 20 questions according to Hess' Cognitive Rigor Matrix. To calculate frequency, the matrix's total number of questions per cell was divided by 20, the total number of questions analyzed, leading to a percentage. Any question coded in cells C,2 through D,6 were determined to have a cognitive complexity of Webb's DOK of Level 3 or 4 and ultimately was considered to have a higher level of cognitive complexity required to demonstrate a logical response as distinguished in Webb's Alignment Tool training manual (Webb, et al., 2005). Any question coded in cells A, 1 through B, 6 was found to have a cognitive complexity of Level 1 or 2 of Webb's DOK and was determined to have a lower level of cognitive complexity. Webb's Alignment Tool (2005) and Hess' Cognitive Rigor Matrix (2009) provided examples as necessary to assist with consistency in the analysis completed by the coders. Hess' Cognitive Rigor Matrix illustrates reliability and validity through the overlapping of Bloom's Revised Taxonomy and Webb's DOK, which have both been found as reliable and valid frameworks. The consistency of the three coders when using the double-rater read behind consensus model illustrates the reliability of cognitive rigor of questions per Hess' Cognitive Rigor Matrix.

Following analysis of the coding of questions, four questions below required a change of cell alignment, which were agreed upon following the double-rater read behind consensus model. To conclude the coding session held on December 6th, 2022 after all four sources of questions were analyzed, the coders reviewed any disparities, which closed with 100% agreement on every question that was discussed.

Chapter Summary

Chapter III described the coding protocol that was utilized to determine the cognitive rigor of 20 questions from the CCRA+ per Hess' Cognitive Rigor Matrix based on research literature. The research questions were answered via a convergent, mixed-methods parallel study. The process for coding, triangulation of data, and ensuring reliability was reached through Mayring's (2000) step model for deductive category application. Mayring's step model provides a visual of the process and methods to reach and ensure credibility throughout the study. Hess' Cognitive Rigor Matrix is a superimposition of Bloom's Revised Taxonomy and Webb's Depth of Knowledge. All three coders were trained in determining Cognitive Rigor based on Webb's Alignment Tool (2005) and Hess' Cognitive Rigor Matrix (2009). A coding protocol was used along with an agenda for clarity and organization of definitions, examples, and rules to increase inter-rater reliability. Before coding, the protocol determined that if consensus was not met, the coders would select the higher level of cognitive complexity per Webb's DOK Alignment Tool (2005). A final coding template of each analyzed question per Hess' Cognitive Rigor Matrix can be found in Appendices B and C.

Chapter IV presents the analytical findings of the study and will answer the overarching and two subquestions that have guided this research study.

Chapter IV

Results

Introduction

Chapter IV presents the findings of the parallel mixed-methods study that examined the type of thinking required by the College and Career Readiness Assessment (CCRA+) for middle and high school students. Twenty retired, publicly available assessment prompts were analyzed for this study. Qualitative content analysis was used to code and categorize the type of thinking required of each question. Descriptive statistics were used to explain the frequency of each type of thinking.

The double-rater read-behind consensus model was used as part of the deductive coding process. The double-rater read-behind consensus model has been found to increase inter-rater reliability effectively (Miles et al., 2014; Sato et al., 2011). Coding training and calibration sessions were conducted. During the calibration sessions, the coders analyzed sample test questions that required various types of thinking.

Elo et al. (2014) wrote, "content analysis processes involve three main phases: preparation, organization, and reporting of results (pp. 1,2)". The preparation component consists of collecting suitable data for content analysis (Elo et al., 2014). The assessment questions that were analyzed can be found on publicly available websites from the Illinois Mathematics and Science Academy (2019), the Wake County Public School System (Council for Aid to Education, 2014), the Virginia Beach City Public School System (Council for Aid to Education, 2018), and from an independent educational consultant via Slideshare (Council for aid to education, 2015)

This study was guided by the overarching question: What type of thinking is required by the College and Career Readiness Assessment (CCRA+) for high school and middle school Students? The research question was followed by two sub-questions that were used for interpreting the data to answer the question both qualitatively and quantitatively.

Findings for Sub-question 1

The first sub-question of this study was: What is the frequency and percentage of questions categorized as higher-level thinking within each cell of Hess's Cognitive Rigor Matrix on the CCRA+?

As suggested by Hess's Cognitive Rigor Matrix, the lowest level of thinking is categorized in cell [A,1], which is comprised of Level 1 thinking on Webb's Depth of Knowledge, and basic recall, remembering, or directly copying information from the provided text or documents. Level 1 thinking is combined with the foundational level of Bloom's revised taxonomy, remembering. Level 2 of Webb's Depth of Knowledge "includes the engagement of some mental processing beyond a habitual response" (Webb, 2007, p. 12). Level 2 questions are not taken from verbatim from a provided document. There is a requirement to demonstrate skills and concepts to answer Level 2 questions accurately. During the calibration of the coders, a determined rule was questions deemed to be found in Level 1 or Level 2 of Webb's Depth of Knowledge would be considered to require lower-level thinking to respond correctly. Any questions coded [A,1], [A,2], [A,3], [A,4], [A,6], [B,2], [B,3], [B,4], [B,6] ultimately were considered lower-level.

Questions and tasks found to align with Level 3 and Level 4 of Webb's Depth of Knowledge were considered to require higher-level cognitive complexity to answer correctly.

Webb (2007) summarized Level 3 as strategic thinking. For a question to prompt Level 3 thinking, a student must "reason, plan, use evidence, and a higher level of thinking than the previous two levels" (p. 6). Level 4 thinking was described as extended thinking and "requires complex reasoning, planning, developing, and thinking" (p. 12). Webb continues to write, "At Level 4, the cognitive demands of the task should be high and the work should be very complex" (p. 13).

Qualitative Findings for Sub-question 1

Twenty questions were analyzed and coded. Seventeen of the questions were selected-response questions, and three questions were performance tasks. There was direct agreement on sixteen of the twenty questions. Indirect agreement was reached on four questions. For any question in which indirect agreement was initially reached, the double-rater read behind consensus model was utilized, and a consensus was reached for direct agreement.

The coders initially disagreed on two of the five Critical Reading and Evaluation selected-response questions from the middle school practice assessment. To answer these questions, students were provided two documents. The first document was a written statement from a school superintendent sent to parents addressing the steps the district is planning to take in response to an increase in crime in school buildings. The second document was a letter written by a parent in the community opposing these actions. Students were prompted to read each document and then answer the corresponding questions. Question 1 of the Critical Reading and Evaluation selected-response section resulted in initial disagreement. The question presented four statements for students to choose from and asked which statement provided the strongest support for the superintendent. Question 2 of the Critical Reading and Evaluation selected-response section also resulted in an initial disagreement. Question 2 asked what can be inferred from four provided responses that the author who wrote to the local newspaper would agree with.

The coders came to consensus that the level of thinking required to accurately answer the question was [B, 2] per Hess's Cognitive Rigor Matrix. The specific language from the Reading-Listening Matrix (2013) that the coders agreed upon was:

- Make basic inferences or logical predictions from data or texts

The coders found indirect agreement when coding Question 3 from the Scientific Quantitative Reasoning selected response section of the middle school practice assessment. The question asked students first to analyze one document with three data sources. The sources included a written statement regarding the number of listeners and the amount of money generated from streaming music. Question 3 provided four answer responses for a student to select from, asking the student to select which statement aligns with the author's claim. The coders came to consensus that this question requires a level of thinking found in cell [B, 2] of Hess's Cognitive Rigor Matrix, based on the specific language:

- Make basic inferences or logical predictions from data or texts

The fourth question the coders initially disagreed on came from a high school student guide. The performance task provided seven documents for students regarding an upcoming mayoral election. The students were tasked to evaluate the claims made by both candidates and make a recommendation endorsing one of the candidates. Following discussion of the thinking required to successfully answer this question, the coders found alignment with cell [D, 5] of Hess's Cognitive Rigor Matrix, based on the specific language:

- Evaluate relevancy, accuracy, & completeness of information from multiple sources
- Apply understanding in a novel way, provide argument or justification for the application

Quantitative Findings for Sub-question 1

Quantitative Findings for CCRA+ assessments that were analyzed in this study are:

- Number and percentage of questions categorized as cell [B, 2] a combination of Level 2 thinking according to Webb's DOK and Understand according to Bloom's Revised Taxonomy - 14 (70%)
- Number and percentage of questions categorized as cell [C,4] a combination of Level 3 thinking according to Webb's DOK and Analysis according to Bloom's Revised Taxonomy - 3 (15%)
- Number and percentage of questions categorized as cell [D,5] a combination of Level 4 thinking according to Webb's DOK and Evaluate according to Bloom's Revised Taxonomy - 1 (5%)
- Number and percentage of questions categorized as cell [D,6] a combination of Level 4 thinking according to Webb's DOK and Create according to Bloom's Revised Taxonomy - 2 (10%)
- Number and percentage of selected-response questions found to require higher-level thinking - 3 of 17 (17.65%)
- Number and percentage of selected-response questions that required students to understand to successfully answer the question - 14 of 17 (82.35%)
- Number and percentage of selected-response questions that required students to analyze to successfully answer the question - 3 of 17 (17.65%)
- Number and percentage of performance task questions found to require higher-level thinking - 3 of 3 (100%)

- Number and percentage of performance task questions found to require students to evaluate - 1 of 3 (33%)
- Number and percentage of performance task questions found to require students to create - 2 of 3 (67%)
- Number and percentage of questions categorized as higher-level thinking, Level 3 or Level 4 according to Webb's DOK - 6 (30%)
- Number and percentage of higher-level thinking questions that required students to analyze - 3 of 6 (50%)
- Number and percentage of higher-level thinking questions that required students to create a response - 2 of 6 (33%)
- Number and percentage of higher-level thinking questions that required students to evaluate in order to respond - 1 of 6 (17%)

Figure 7 depicts a visual of the percentage of higher level thinking questions overall, on selected-response questions, and on the performance tasks. Overall, 30% of the questions required higher level thinking. These questions required students to either analyze, evaluate, or create in order to successfully answer a question. The three selected-response questions that required higher-level thinking required students to analyze information in order to create a correct answer. These three questions coded in cell [C, 4]. Of the three performance tasks, two of the assessment prompts required students to create an answer based on the provided documentation, cell [D, 6]. The third assessment prompt, required students to evaluate a claim that was made based on provided documents cell [D, 5].

The national average of higher-level questions asked during classroom lessons is approximately 15% (Tienken, Goldberg, DiRocco, 2009). The percentage of questions categorized as higher level on the CCRA+ exam was double the national average when determining the average number of higher-level questions asked during a classroom lesson. The type of thinking, however, is highly repetitive.

Figure 7

Cognitive Complexity of All Analyzed Questions

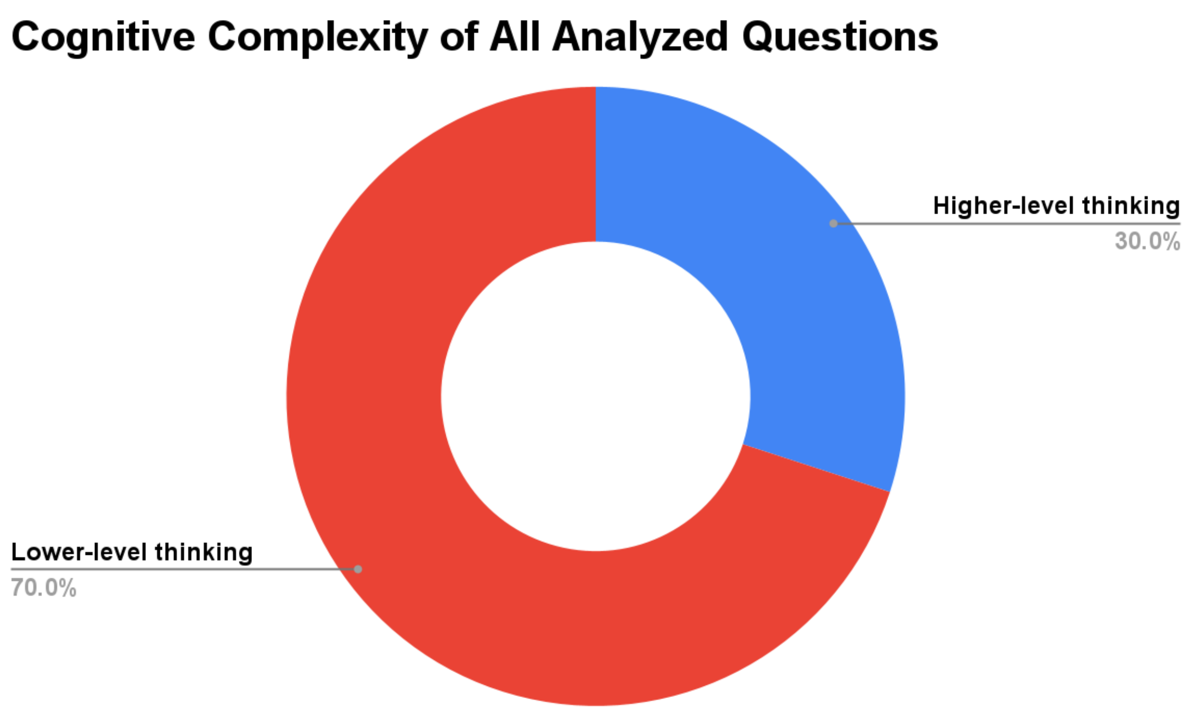
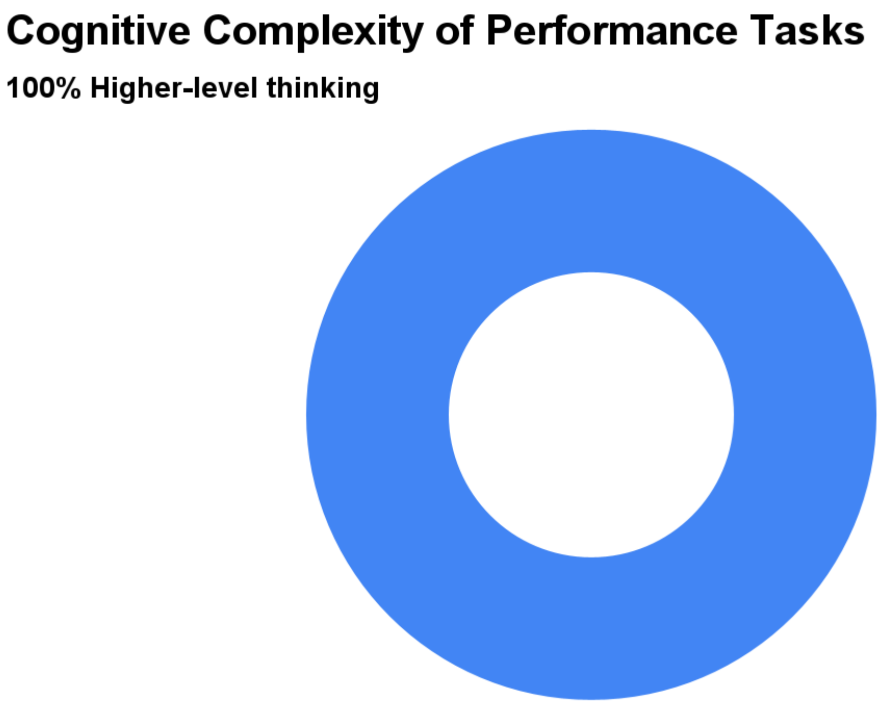


Figure 8 depicts the cognitive complexity of the performance tasks that were analyzed by the coders. All of the performance tasks, 100%, required higher level thinking.

Figure 8

Cognitive Complexity of Performance Tasks



Bloom’s Revised Taxonomy measures the type of thinking required to answer a question correctly. The seventeen selected-response questions only required two of six types of thinking. Fourteen of the seventeen questions required thinking aligned to the understanding level of Bloom’s Revised Taxonomy, the remaining three selected-response questions corresponded with the level of analyze on Bloom’s Revised Taxonomy. Two of the three performance tasks that were analyzed required thinking that aligned with the create level of Bloom’s Revised Taxonomy, the third performance task question required thinking that required thinking in line with the level of evaluate.

Table 3

The Type of Thinking Required for Each Analyzed Question based on Bloom’s Revised Taxonomy (2001)

	Selected-Response Questions		Performance Tasks	
	Frequency	Percentage	Frequency	Percentage
Remember	0	0	0	0
Understand	14/17	82.35%	0	0
Apply	0	0	0	0
Analyze	3/17	17.65%	0	0
Evaluate	0	0	1/3	33%
Create	0	0	2/3	67%

Findings for Sub-question 2

The second sub-question of the study was: In what ways are the skills identified as essential for success in the workplace by global organizations assessed in the CCRA+? Bloom's Revised Taxonomy (2001) is the framework most suitable to determine the type of thinking required to demonstrate the skills essential for success in the workplace. Understanding content through Bloom's Revised Taxonomy is not linear, however the type of thinking demonstrates the level of content mastery or skill. Qualitative Findings for Sub-question 2

The WEF (2020) suggested skills such as "complex problem-solving," "critical thinking and analysis," and "creativity, originality and initiative" are of greatest need to be successful in the workforce. Parallel to the skills identified as essential, the Council for Aid to Education (CAE, n.d.) provides a suite of assessments that claim to measure "college and career readiness skills of critical thinking, problem-solving, and effective written communication." The skills that CAE claim to assess through the CCRA+ as well as the skills the WEF have identified as essential for success in a career align with the language found in assessments and thinking that would be considered Level 3 or Level 4 based on Webb's Depth of Knowledge.

Critical thinking and originality are foundational components in order to successfully answer a Level 3 or Level 4 question according to Webb's Depth of Knowledge. Examples of higher-level thinking from questions on the CCRA+ performance task according Hess's Cognitive Rigor Matrix for Writing-Speaking (2013) are:

- Develop and explain opposing perspectives or connect ideas, principles, or concepts using supporting evidence (quote, example, text reference, etc.)
- Develop arguments of fact (e.g. Are these criticisms supports by the historical facts? Is this claim or question true?)
- Use reasoning and evidence to generate criteria for making and supporting an argument of judgment (Was FDR a great president? Who was the greatest ball player?)
- Support conclusions with evidence
- Evaluate validity and relevance of evidence used to develop an argument or support a perspective
- Describe, compare, and contrast solution methods
- Develop an alternative solution or perspective to one proposed (e.g. debate)
- Use multiple sources to elaborate on how concepts or ideas specifically draw from other content domains or differing concepts (e.g. research paper, arguments of policy - should this law be passed? What will be the impact of this change?)
- Compare and contrast conflicting judgments or polices (e.g., Supreme court decisions)
- Synthesize information across multiple sources or texts in order to articulate anew voice, alternate theme, new knowledge or nuanced perspective

Similarly, of the selected-response questions that the coders found to require higher-level thinking, to select the correct answer, the following skill was needed: Analyze information within data sets or texts

To successfully answer the twenty assessment prompts analyzed, students have to demonstrate the ability to think in four different ways according to Bloom's Revised Taxonomy. The frequency of the four types of thinking varied greatly. The most frequent type of thinking was the lowest type of thinking found on the test. Fourteen of the seventeen selected-response questions required students to understand according to Bloom's Revised Taxonomy (2001). The fourteen questions all required the student to infer Anderson et. al, (2001) defined inferring as, "finding a pattern within a series of examples or instances…a student is able to abstract a concept or principle that accounts for a set of examples or instances by encoding the relevant features … and, most important, by noting relationships" (p. 74). The fourteen questions coded to require understanding as the type of thinking is not the type of thinking required to be successful in the workforce as this type of thinking is easily replicated by artificial intelligence, therefore not critical thinking or problem solving. Three of the selected-response questions required students to analyze information in order to successfully answer the question. Analysis, per Bloom's Revised Taxonomy is defined as "breaking material into its constituent parts and determining how the parts are related to one another and to an overall structure" (Anderson, et al., 2001, p. 109). In order to successfully answer the performance task questions, students were required to use flexible thinking. Students would need to vary their type of thinking throughout Bloom's Revised Taxonomy in order to create a coherent response to the assessment prompt, inclusive of understanding the material provided in the prompt, applying the material accurately, and ultimately evaluating or creating a response to what was asked of the student.

Quantitative Findings for Sub-question 2

Table 4 presents the total frequency of questions categorized based on Webb's DOK Levels and overall percentages. The table illustrates that none of the questions coded on the CCRA+ were Level 1 of Webb's Depth of Knowledge. The majority of selected-response questions were found to be Level 2 questions. All performance task questions required Level 4 thinking to successfully answer. Questions must require Level 3 or Level 4 thinking to be considered to accurately assess the skills identified as essential for success in the workplace by global organizations. Table 4 illustrates the findings of level of thinking required to successfully answer selected-response questions and performance tasks based on higher-order thinking skills from Bloom's Revised Taxonomy (2001). Bloom's Revised Taxonomy is comprised of six levels, according to Anderson et al., (2001):

> Objectives that involve problem solving and critical thinking most likely call for cognitive processes in several categories on the process dimensions. For Example, to think critically about an issue probably involves some *Conceptual Knowledge to Analyze* the issue. Then, one can *Evaluate* different perspectives in terms of the criteria and, perhaps. *Create* a novel (pp. 269-270)

The three levels of Bloom's Revised Taxonomy (2001) considered to require higher order thinking are analyze, evaluate, and create. The frequency column indicates the volume of questions that were coded to require the aligned level of thinking, the percentage column is found by calculating the frequency.

Table 4

Coding Chart of All Analyzed Questions based on Bloom's Revised Taxonomy (2001)

	Selected-Response Questions	Performance Tasks

	Frequency	Percentage	Frequency	Percentage
Does not require higher order thinking	14/17	82.35%	0	0
Does require higher order types of thinking	3/17	17.65%	3/3	100%

Conclusion

The purpose of this convergent, parallel mixed-methods study was to analyze and describe the cognitive complexity of the questions on the CCRA+, published by the CAE, compared to the language of higher-level thinking found in the research literature. Hess's Cognitive Rigor Matrix was the framework used to determine the cognitive of complexity and categorize each question based on the language used in the assessment questions. Any question that was coded as Level 1 or Level 2 of Webb's Depth of Knowledge was considered to only require lower-level thinking and ultimately will not determine student preparedness for success in either college or their career. Questions or tasks that were coded to require thinking that aligned with Level 3 or Level 4 on Webb's Depth of Knowledge were considered to require critical thinking that would illustrate preparation for success in the workforce. Bloom's Revised Taxonomy (2001), is a framework used to identify the type of thinking students use to answer an assessment prompt. Assessment prompts that were considered lower-level thinking aligned with the categories remember, understand, or apply. Higher-level thinking aligned with the categories analyze, evaluate, and create.

Chapter IV presented results related to the convergent, parallel mixed-methods study to describe the type of thinking that is required by the College and Career Readiness Assessment (CCRA+). The findings were reported through an analysis of the frequency of questions categorized as requiring higher-level thinking and describing the ways the skills that are identified as essential for success in the workplace by global organizations are assessed and measured.. In most instances, requiring students to explain their aligns with Level 3 thinking per Webb's Depth of Knowledge (Webb, 2007, p. 6). "Students should be required to make several connections – relate ideas within the content area, or among content areas – and would have to select one approach among many alternatives on how the situation should be solved" (Webb, 2007, p. 13). The coding of questions revealed the performance tasks on the CCRA+ for both middle school and high school assessments, inclusive of document analysis and a written, constructed response required the higher level of thinking that is identified in Hess's Cognitive Rigor Matrix.

Four of the types of thinking from Bloom's revised taxonomy were represented in the analyzed assessment prompts. The frequency of the questions per type of thinking greatly varied. Not only did the selected-response questions require lower-level thinking per Webb's Depth of Knowledge, students were only required to think to the level of understand over 82% of the time. The assessment prompt asked students to make an inference from the information provided in the document resources. Students did not have to think at a higher-level or vary the type of thinking to correctly answer a question.

Chapter V presents a summary of the methodology of the study and discussion of the findings related to the two research subquestions. Chapter V continues with implications for policy and practice for both school leaders and policy makers. Future recommendation for continued research will close Chapter V.

Chapter V
Conclusion

Chapter V presents conclusions, recommendations for policy and practice, and recommendations for future research. This parallel mixed-methods study aimed to describe the type of thinking required by retired, publicly released items from the College and Career Readiness Assessment for middle and high school students. The study was guided by the overarching research question: What type of thinking is required by the College and Career Readiness Assessment (CCRA+) for middle and high school students?

Two sub-questions guided the study:

a.) What is the frequency and percentage of questions categorized as higher-level within each cell of Hess' Cognitive Rigor Matrix on the CCRA+?

b.) In what ways are the skills identified as essential for success in the workplace by global organizations assessed in the CCRA+?

Conclusions

Creating selected-response questions, for standardized tests that require higher-level thinking is difficult. Previous studies have shown the thinking requirements found on nationally used selected-response, standardized assessments and remediation programs are staggeringly lower-level thinking questions. Solis-Stovall (2020) found 90% of the questions on grade 3 and grade 4 PARCC assessments required lower-level thinking to answer correctly (p. 89). Lamberti (2020) found 96% of questions found on the TerraNova Grade 8 assessment and 100% of questions found on the IOWA Grade 8 assessment required Level 1 or 2 thinking according to Webb's Depth of Knowledge (pp. 104-105). Dorrian (2021) found 94.85% of tenth-grade

PARCC assessment questions required lower-level thinking. Sydoruk (2018) found 88%, of questions on an online remediation program that claimed to analyze higher-level thinking actually required Level 1 or Level 2 thinking per Webb's DOK (p. 100).

The results from this study support the writing of Dewey (1910) about the limitations of instruction based on predetermined answers: "If the suggestion that occurs is at once accepted, we have uncritical thinking" (p. 13). Selected-response questions provide students with suggested answer options to choose from. The selected-response question type allows students to answer quicker than any other format as the responses are provided. However, this format rarely offers the opportunity to think critically or strategically and does not allow a student to extend their thinking. Selected-response assessments are the most efficient question type in terms of speed required for both student completion and grading. Dewey (1910) warned of a focus on efficiency:

> In the mastery of reading, writing, drawing, laboratory technique, etc., the need of economy of time and material, of neatness and accuracy, of promptness and uniformity, is so great that these things tend to become ends in themselves, irrespective of their influence upon general mental attitude. Sheer imitation, dictation of steps to be taken, mechanical drill, may give results most quickly and yet strengthen traits likely to be fatal to reflective power. (p. 58)

The peril of assessments that focus on efficiency lies in the focus on lower thinking they tend to entrench in teaching and learning. Original thought is not facilitated with selected-response questions. Student engagement is limited because, on tests populated with a preponderance of selected-response items, higher-level thinking is usually not required to select a pre-determined correct answer. The thinking skills used to answer selected-response questions reinforce rote thinking. Rote thinking has been replaced throughout the knowledge economy and is completed by bots or other forms of artificial intelligence (AI).

Questions in which the answer is provided for the student can lead to functional fixedness (Runco and Chand, 1995). Runco and Chand (1995) explained “as individuals who rely on cues tend to be unoriginal, so too can individuals who rely on details from previous experience- even persons who have been innovative in the past- preclude original behavior” (p. 248). With the majority of questions on standardized assessments being selected-response, teachers spend time on test-taking strategies in place of knowledge acquisition or investing time in critical thinking activities.

When students are provided an answer to choose from, the need for students to use higher order thinking decreased and education becomes the pursuit of finding one correct answer. Aikin (1942) commented, “our secondary schools did not prepare adequately for the responsibilities of community life. Schools generally were excellent examples of autocratic, rather than democratic, organization and living”. (p. 4) Through a teaching style in which students are provided a potential answer, students become focused on compliance, identifying themes, and learning the personality of the teacher rather than generating knowledge from a yearning of curiosity. Dewey (1910) contrasted information and wisdom, which described students who learn from an autocratic style as those who seek information or lower-level thinking. Students who seek wisdom are those with the ability to transfer knowledge and go beyond previous lessons and recall:

> Information is knowledge which is merely acquired and stored up; wisdom is knowledge operating in the direction of powers to the better living of life. Information, merely as information, implies no special training of intellectual capacity; wisdom is the finest fruit of that training. In school, amassing information always tends to escape from the ideal of wisdom or good judgment. (p. 59)

Frequent use of teaching techniques that require students to find correct answers and regurgitate them, coupled with doses of standardized tests that consist mainly of selected-response questions, can cause functional fixedness. Wiggins (2011) identified the discrepancy between students who perform well on standardized assessments yet struggle in authentic environments:

> The often-strange quality of new knowledge can cause us to unwittingly misunderstand new ideas by assimilating them into our old conceptions…That is why so many students who do so well on school tests seem so thoughtless and incompetent in solving real-world problems. (p. 84)

Teachers must provide and guide students to develop original, creative thought that is not demonstrated by answering lower-level thinking, selected-response questions.

Seifert (2009) explained, “Over time, we get so used to one particular purpose for an object that we overlook other uses” (p. 190). When students engage with one type of thinking repeatedly for answering assessment prompts, expecting a shift in the type of thinking to answer the few randomly inserted higher-level thinking questions is unlikely. Students are not accustomed to changing their type of thinking to answer selected-response questions.

The CRAA+ provides an effective assessment of how format can drive higher level thinking according to Bloom’s Revised Taxonomy and Hess’s Cognitive Rigor Matrix. All three performance task questions on the CRAA+ were found to require Level 4 thinking on Webb’s Depth of Knowledge. There is a possibility CRAA+ provides an example to design a standardized assessment to gauge students’ ability to think critically, problem solve and ultimately be prepared for college or a career beyond high school graduation. Through analyzing multiple informational sources, synthesizing documents, and creating original products, students are required to think at the highest level on the CRAA+ performance tasks. As Dewey (1910) wrote, “Judgment reveals the bearing or significance of facts: Synthesis. As analysis is conceived to be a sort of picking to pieces, so synthesis is thought to be a sort of physical piecing together”

(p. 135). The performance task assessment prompts on the CCRA+ required students to analyze, synthesize, evaluate, and create meaning through the use of higher-level thinking.

When a student is given a task to create or evaluate, they must first have an understanding of the information, be able to apply the information in a coherent way to communicate their response, and potentially analyze different aspects or views. Assessment prompts that require the highest level of thinking unintentionally require students to be able to transfer knowledge and switch their type of thinking. Seifert (2009) wrote of a common dilemma of functional fixedness, "the tendency for a person to frame or think about each problem in a series in the same way as the previous problem, even when doing so is not appropriate to later problems" (p. 190). The frequent use of different types of thinking for different purposes helps defend against functional fixedness.

The findings of this mixed-methods parallel study suggest that both the level of thinking, as determined through Webb's Depth of Knowledge (1997) and the type of thinking, as determined through Bloom's Revised Taxonomy (2001) are essential to create students who can demonstrate flexible thinking. Only through the ability to transfer knowledge in different formats will students be able to meet the skills necessary to be successful in the workforce as determined by leading organizations, such as the World Economic Foundation, IBM, Dell, Intel, etc.

The findings of this study indicate that the format of an assessment question matters in that the format drives the function of thinking. When the same type of thinking is required by high stakes tests, teachers develop lessons in response that align with the format of the tests (Au, 2007). If an expectation is for students on high stakes tests is to find correct answers to selected-response items, then teachers will use classroom instructional strategies and activities that require students to memorize facts and find correct answers (Au, 2007).

If the format of a high-stakes assessment requires students to demonstrate the ability to use several types of thinking to evaluate a topic or create an original response, the classroom instruction and activities will include such thinking (Tienken, 2017). Once authentic assessments are the expectation, curricula development, unit development, lesson planning, objective creation, and ultimately instruction will follow in nature. Authentic assessments will lead to authentic instruction, as exposure to this type of learning will be the only way students can demonstrate success on their performance tasks.

Through authentic lessons and assessments, students will have more opportunities to practice and demonstrate flexible thinking according to Bloom's Revised Taxonomy (2001) and higher-level thinking per Webb's Depth of Knowledge (1997). Only after curricula development, instructional practice, assessment design, and feedback to students are authentic can educators truly expect students to meet the demands of becoming problem solvers who think creatively and critically.

The existing literature, theory, and results of this study suggest there is an obligation on the part of policymakers and educators to rethink the types of standardized tests currently in use at the state level. Knowing standardized assessments have the ability to provide prompts that require higher-level thinking, educators, school administrators, and policymakers bear the responsibility of selecting and administering assessments that pose questions found to require higher-level cognitive complexity.

Assessment is the culminating activity of acquired knowledge, skill, or thinking. To demonstrate flexible, higher-level thinking on an assessment, students must be engaged in active learning experiences that require students to think across the six levels of Bloom's Revised Taxonomy (2001) as well as the different levels of Webb's Depth of Knowledge (1997). If

students are not required to think actively through authentic, community centered, problem and project-based learning regularly in classroom lessons, projects, and units, the expectation for students to demonstrate these skills on an assessment is misaligned and unrealistic.
Several educational shifts must be made to adequately prepare students to be successful in the workforce.

Recommendations for Practice

Many teachers and administrators are familiar with cognitive complexity frameworks such as Webb's Depth of Knowledge and Bloom's Revised Taxonomy, yet higher level thinking activities and assessments are under-represented in the public education environment, with the majority of questions encountered by students requiring low level thinking (Tienken, Goldberg, DiRocco, 2009). With pressure from multiple sources, educators can succumb to the allure of standardized test preparation, teaching to the test, providing simulations of the current state-mandated standardized tests, and participating in practices that do not allow students to develop higher-level thinking skills.

Higher level thinking and cognitive complexity must be a priority of classroom instruction and focus of the professional development of all staff members. Curricula, planning, lesson objectives, instruction, and ultimately assessment must represent authentic life situations to engage students in active learning and higher level thinking (Dewey, 1910). Through active learning, students will engage in the cognitive skills that are expected to be successful in the workforce. Students can only demonstrate mastery and hone critical thinking when they are given an environment where higher level thinking is regularly facilitated and practiced. Authentic education must become the core of the practice, which will then prepare students for success on assessments due to fostering a variety of types and depths of thinking.

School leaders should provide teachers with support and coaching to develop sustainable curricula that include objectives and activities aimed at developing higher-level thinking. A responsibility of education is to "cultivate deep-seated and effective habits of discriminating tested beliefs from mere assertions, guesses, and opinions; to develop a lively, sincere, and open-minded preference for conclusions that are properly grounded, and to ingrain into the individual's working habits methods of inquiry and reasoning" (Dewey, 1910, p. 32). The habits Dewey (1910) referenced are developed through actively exploring the problems of society. When teachers are adequately trained to create active learning environments for students, and administrators are equipped to support their staff with suggestions and enhancements, students can be provided opportunities to deepen their ability to think and act in cognitively complex ways.

School leaders should revise curricula to include higher-level project-based learning (PBL) activities to expose students to authentic activities that mirror the type of thinking required for success beyond high school graduation. According to Tienken (2020) "Learning does not get more active or authentic than a socially conscious PBL based on student interests and passions" (p. 123). Through project-based learning, curricula helps students engage in the fundamental necessities highlighted by Dewey (1910, 1916), The Cardinal Principles of Secondary Education (1918), and The Eight-Year Study (1942). Project-based learning offers the potential to engage students in local topics in a way that fosters a connection with the community and individual identity. Tienken (2020) writes:

> The selection of a problem or issue based on careful consideration of what is already occurring in society addresses a pressing social force and aligns with another important aspect of the Curriculum Paradigm: the nature of the learner as an active constructor of meaning. (p. 123)

Project-Based Learning must culminate with authentic assessments that require some higher-level thinking opportunities. School districts are recommended to develop curricula that allows time for authentic assessments. Wiggins (2011) writes, “Authentic tests do not rely on unrealistic and arbitrary time constraints, nor do they rely on secret questions or tasks” (p. 90). This practice is a shift from current standardized assessments that are a one size fits all test with a constricted amount of time that does not demonstrate flexibility for the student to go into greater depths or demonstrate learning at a pace corresponding with their cognitive development.

After ensuring school curricula includes project-based learning, teachers must be supported to practice with autonomy, freedom and creativity, knowing all students will not complete assessments or units at the same pace. Professional development will focus on creating authentic assessments and how to differentiate lessons for students to provide supplemental resources for students to stay engaged in all components of their authentic learning experience while ensuring opportunities for higher-level thinking are provided for students.

Schools must eliminate the use of lower-level thinking assessments that include selected-response questions. The selected-response assessment format is inadequate and outdated. The current assessment practice of schools allows the possibility “for a student to pass all of his or her courses and still remain functionally and culturally illiterate” (Wiggins, 2011, p. 82). School district personnel can help prepare students for success beyond graduation with an on-going review of curricula through a critical lens to ensure deep and flexible thinking is incorporated through socially relevant lessons in the classroom.

Curriculum development must include an analysis of types of thinking and depth of thinking to ensure proper scaffolding throughout the curriculum while eliminating the unintentional lure of creating learning experiences that indirectly lead to functional fixedness. Creating an authentic curriculum that aligns with the experiences of the students in the learning community will reduce the need for educators to rely on packaged curriculum programs that focus on lower level thinking.

Only through active learning and authentic assessments will students be required to think in different formats and at high levels. Homegrown curricula is the only way to ensure the community and first hand experiences of students are incorporated into what is being taught and assessed. As Dewey, (1910) wrote:

> The instructor ceases and the teacher begins at the point where communicated matter stimulates into fuller and more significant life that which has entered by the straight and narrow gate of sense-perception and motor activity. Genuine communication involves contagion; its name should not be taken in vain by terming communication that which produces no community of thought and purpose between the child and the race of which he is the heir. (p. 271)

Educators must become well versed in Hess's Cognitive Rigor Matrix to help them differentiate between depth of thinking and type of thinking. Educators should have opportunities to expand their instructional craft through peer observations and professional development that focuses on developing and strengthening a community partnership and authentic learning experiences.

Assessments should be redesigned to require the demonstration of multiple types of thinking as well as a deep level of thinking. Formative assessment will always remain as a quick, efficient tool to ensure that students are grasping material, however, summative assessments should include portfolios and presentations, practices that reflect critical-thinking while allowing students to demonstrate their problem-solving abilities.

Finally, curricula revisions must be on-going and a practice that includes all educators. Providing educators with time to reflect on and revise curriculum based on student needs will lead to a document that is followed with a greater level of fidelity and accuracy. Curriculum should include partnerships and activities that are inclusive of the students' community as this is an area students are familiar with and likely to create deep connections to and meaningful learning with. The cycle of providing professional development, revising curricula, implementing active learning in student activities, and assessing authentically will lead to understanding students preparation for their time beyond high school.

Recommendations for Policy

The fourth industrial revolution marks the rapidly expanding use of AI to automate lower-level thinking occupations. Educators must prepare students to complete tasks that are beyond the cognitive capabilities of artificially intelligent machines and apparatuses or that complement the abilities of AI. AI programs that are capable of writing multiple-paragraph responses, create poetry, develop screenplays, and design lecture presentations are now publicly available, free, and highly accurate. The ability for students to be able to analyze information and produce responses in a variety of formats that fit the appropriate audience is necessary. Without the ability to problem-solve and think creatively, the thinking of individuals will be replaced by machines.

AI chatbots have disrupted education with ease of use and accuracy to detail along with skillful writing. Educators are facing a divide in which this software can be integrated into education propelling higher-order thinking or can be blocked. If the decision of school districts through policy is to exclude AI practices will continue to increase the lower-level thinking that is prevalent in the current state standards and state standardized assessments. Building in the use of AI will prepare students with the use of tools that will be available for use beyond the classroom and in their postsecondary careers. Education policy can help school district personnel to develop the necessary skills required of students for success in the workforce through instructional practices built into curriculum and daily learning activities of students.

State policy, with input from university preparation programs and educators, needs to be reviewed and revised, beginning with teacher preparation standards. Landmark studies such as the Eight-Year Study (1942), the Cardinal Principles of Secondary Education (1918), and the works of Dewey (1910, 1916) provide results and exemplars of how districts implemented curricula and opportunities for students to learn and develop higher-order thinking skills actively. All teacher preparation programs should be free to include coursework that presents educational history and highlights of programs that have proven to be effective. Through this knowledge, teachers in training will develop an understanding of the activities and practices that have previously been shown to develop skills students need to be successful in life. With this knowledge taught in preparation programs, teachers will be equipped to provide authentic learning experiences for their students that can connect with their local community and develop deeper knowledge.

As Tienken (2018), wrote "Standardized test results do not capture accurately what or how well students learn, especially when students are subjected to lard doses of test preparation" (p. 57). States should replace standardized tests as the arbiter of success on school performance reports with a portfolio approach. One piece of that portfolio could include student presentations about problem-based learning assignments in which students use their knowledge and personal interests and experiences to propose solutions to problems within society. As Dewey (1910) wrote:

> Thinking is not like a sausage machine which reduces all materials indifferently to one marketable commodity, but is a power of following up and linking together the specific suggestions that specific things arouse. What geometry does for one, the manipulation of

> laboratory apparatus, the mastery of a musical composition, or the conduct of a business affair, may do for another. (p. 55)

Student presentations will vary depending on the interests of students. The variation in the problems studied and solved by students' mimics how members of organizations provide different strengths and capabilities to a situation. Through multiple presentations, students will have a body of work compiled in the form of a cumulative folder or transcript which illustrates the growth of the student. As written in The Cardinal Principles of Secondary Education (1918), "Education must be conceived as a process of growth. Only when so conceived and so conducted can it become a preparation for life" (p. 16). An amended form of a student's transcript, similar to that used by the Mastery Transcript Consortium should be an option for New Jersey public schools.

Examples of successful portfolio based assessment models are in existence. The New York Performance Standards Consortium is a group of 38 public schools that have developed authentic assessments that gauge mathematical problem solving, technology use, science research, service learning, career skills, and visual and performing arts. (Tienken 2018), From 2000 through 2009, Nebraska used the School-based Teacher led Assessment and Reporting System (STARS). As suggested by Tienken (2018), "the framework, including state policy documents, assessments, and protocols still exist; and state education leaders could easily reinvigorate the system" (p. 58). Providing school districts local control of the assessments that will fulfill the graduation requirement falls in line with the tenets written by Dewey in *How We Think (1910)*.

The final policy recommendation is a revision of the state standards. The current New Jersey Student Learning Standards have a focus on breadth rather than depth and do not allow educators the ability to provide the time needed to develop higher-level thinking skills in a productive manner. Tienken (2017) writes, “standardization of curriculum expectations leads to subject-centered instruction and it results in a system that seeks to force children to conform instead of a system built on the premises of developing creativity and meeting the learning needs of the child” (p. 5). Reducing the number of standards will allow for curriculum development and standards that fit into the local school districts allowing for an experience that is more likely to fit students precisely. As written by Tienken (2017), “curricula designed, developed, and implemented in the twenty-first century must be diversified, pliable, less standardized, and connected to the unique needs and contexts of the students compelled to experience them” (p. 106).

Curricula can only be created for the students in a specific district and implemented with fidelity if there is a reduction in the standardization of public school education. Standards are written in a way that leads to lower-level thinking assessment questions. With the current standards being written in a way that can be assessed through selected-response questions, content only needs to be delivered in a rote manner, as questions only require memorization or understanding to successfully answer, not student creation, a transfer of knowledge, or a blend of types of thinking. Policymakers should view standard creation through a different conception. By looking at knowledge through a different lens, we will broaden what is taught to students. The current standards teach students in a narrow way that does not promote the transfer of thinking, applying what has been taught in different ways, or meaningful cross curricular activities. Standards should broaden student learning and open the thinking of students to allow for wider and accurate application. Standards should promote sustainable learning activities across content areas, breaking down the silos of education that dominate the current structure of secondary schools. The change of standards is the first step in a reduction of the amount of, standardized assessment days in school calendars and an increase in authentic activities. Only through authentic experiences and active learning will students be prepared with the mental skillset to succeed in a college or career.

Recommendations for Further Research

Further research is required to provide an educational landscape that demonstrates higher-order thinking of students as the expectation. In order to create this expectation, student assessments and pedagogical opportunities must align with cognitive development to achieve higher-level thinking. This mixed-methods study only coded twenty questions from retired, publicly available CCRA+ exams. Through additional analysis, themes and question creation can potentially be found that are inclusive of greater percentages of higher-level thinking questions. The creators of the CCRA+, Council for Aid to Education (CAE), provide a suite of assessments that claim to assess student's preparation for college or a career. Further research is needed on the suite of assessments created by CAE. Organizations other than CAE make comparable claims and attempt to gauge student's preparedness for success beyond high school, further research needs to be completed on these assessments as well. Through further analysis, there is potential for an assessment to be created which combines high-level thinking questions in an authentic format throughout the entirety of the assessment.

Additional analysis of student preparedness for college or career of students based on their performance on the CCRA+ should be completed in a longitudinal study. Through the analysis of a cohort of students who participated in the CCRA+ and the outcomes in their postsecondary endeavors, validity can be strengthened in the CCRA+ individual student reports. This research can effectively determine success through achievement in a postsecondary institution or immediately in a career.

Further research should analyze trends of student performance on standardized assessment when authentic assessment is included and completed with fidelity in district curricula. This research will indicate if the claim of authentic lessons that promote active learning leads to a transfer of knowledge that will indicate success on standardized assessments. In addition to identifying authentic assessments, other aspects of lessons and teaching structure that have aligned with higher levels of cognitive thinking should be researched to determine the outcome of student performance on standardized assessments and student's ability to transfer knowledge.

Further research will assist in determining the postsecondary skills that are developed through an internship program. An evaluation of the use of community resources and provided experiences for students as implemented in curricula can move forward the body of research that determines best practices in developing students who possess the skills of critical thinking, problem-solving, and creativity through connections to their community and previously established knowledge a claim made by Dewey in *How We Think* (1910). Through providing authentic internship experiences, further research will be able to determine if the development of soft skills such as collaboration, communication, decision making, and time management transfer to positive results in an academic setting or institute.

Finally, additional research should be conducted in regard to the use and implementation of Hess's Cognitive Rigor Matrix. This research will provide school districts with a model to use a cognitive framework that superimposes two widely used models, Webb's Depth of Knowledge and Bloom's Revised Taxonomy, separately. Through the utilization of Hess's Cognitive Rigor Matrix, teachers and administrators will ensure students are granted learning opportunities that lead to the development of higher-level thinking skills, and an ability to create through a depth of thinking. At the same time, students will be given the opportunity to demonstrate flexible thinking through identifying the impact of teaching to all levels of Bloom's Revised Taxonomy.

References

ACT. (2019). *The condition of college & career readiness 2019*. ACT, Inc. https://www.act.org/content/act/en/research/reports/act-publications/condition-of-college-and-career-readiness-2019.html

Adams, N. E. (2015). Bloom's taxonomy of cognitive learning objectives. *Journal of the Medical Library Association: JMLA*, *103*(3), 152–153. https://doi.org/10.3163/1536-5050.103.3.010

Aikin, W. (1942) *The story of the eight-year study with conclusions and recommendations*. (Vol. 1). Harper & Brothers.

Airasian, P., Cruikshank, K., Mayer, R., Pintrich, P., Raths, J., Wittrock, M., Anderson, L. W., & Krathwohl, D. R. (2001). *A taxonomy for learning, teaching, and assessing*. Longman.

American Psychological Association. (2010). *Using the global assessment of functioning scale to demonstrate the importance of inter-rater reliability*. https://www.apa.org/ed/precollege/undergrad/ptacc/inter-rater-reliability

American Psychological Association. (2022). *APA dictionary of psychology.* American Psychological Association.

Armstrong, P. (2010). *Bloom's taxonomy*. Vanderbilt University Center for Teaching. https://cft.vanderbilt.edu/guides-sub-pages/blooms-taxonomy/

Au, W. (2007). High-stakes testing and Curricular Control: A qualitative metasynthesis. *Educational Researcher*, *36*(5), 258–267. https://doi.org/10.3102/0013189x07306523

Bloom, B. S. (Ed.). Engelhart, M. D., Furst, E. J., Hill, & W. H., Krathwohl, D. R. (1956). *Taxonomy of educational objectives, Handbook I: The cognitive domain*. David McKay.

Boote, D. N., & Beile, P. (2005). Scholars before researchers: on the centrality of the dissertation literature review in research preparation. *Educational Researcher*, *34*(6), 3–15. https://doi.org/10.3102/0013189X034006003

Bracey, G. W. (2008, June) "Cut scores, NAEP achievement levels and their discontents" *School Administrator*, 65, 6: 20–23. https://eric.ed.gov/?q=Advanced+Placement+OR+AP+U.S.+History&=id&ff1=autBracey%2c+Gerald+W.&id=EJ797406

CAE and Johns Hopkins University collaborate on college and career readiness research project—Council for Aid to Education. (n.d.). https://cae.org/cae-and-johns-hopkins-collaborate-on-college-and-career-readiness-project/

CAE named 2021 tech Edvocate award finalist for best learning assessment or tool. (2021, September 16). GlobeNewswire News Room.https://www.globenewswire.com/news-release/2021/09/16/2298451/0/en/CAE-Named-2021-Tech-Edvocate-Award-Finalist-for-Best-Learning-Assessment-or-Tool.html

Cann, O. (2016, September 27). *What is competitiveness?* World Economic Forum. https://www.weforum.org/agenda/2016/09/what-is-competitiveness/

Council for Aid to Education. (n.d.). *Ensure students have the essential skills for academic and career success*. https://cae.org/grades-6-12/

Council for Aid to Education. (2018). *CWRA Student Guide Institution*. Student Guide. https://cdnsm5-

ss19.sharpschool.com/UserFiles/Servers/Server_78010/File/Students/Academics/High%20School/College%20and%20Work%20Readiness/CWRA_Student_Guide_Institution.pdf

Council for Aid to Education. (2014). *CWRA Plus Practice Performance Task*. High School Sample Performance Task. http://wcpssenglish.pbworks.com/w/file/fetch/100357381/CWRA_Plus_Practice_PT.pdf

Council for Aid to Education. (2015). *Middle School Sample Assessment*. CWRA Middle School Practice Assessment. https://www.slideshare.net/ZoeOBrian/07cwramiddleschoolpracticeassessment

Department of the Interior, Bureau of Education. (1918). *Cardinal principles of secondary education* (Bulletin 1918, No. 35). Washington Government Printing Office.

Dewey, J. (1902). *The child and the curriculum.* Chicago, IL: University of Chicago Press.

Dewey, J. (1910). *How we think.* D C Heath. https://doi.org/10.1037/10903-000

Dorrian, H. (2021). *Analysis of the higher order thinking requirements of PARCC practice assessments in grades 10*" [Doctoral dissertation, Seton Hall University]. Seton Hall University Dissertations and Theses ETDs. https://scholarship.shu.edu/dissertations/2876/

Elo, S., Kääriäinen, M., Kanste, O., Pölkki, T., Utriainen, K., & Kyngäs, H. (2014). Qualitative content analysis. *SAGE Journals*, *4*(1), 215824401452263. https://doi.org/10.1177/2158244014522633

Ennis, R. H. (2018). Critical thinking across the curriculum: A vision. *Topoi, 37*(1), 165–184. https://doi.org/10.1007/s11245-016-9401-4

Encyclopedia Britannica, Inc. (2002). Functional fixedness in p*sychology*. In J.E. Luebering (ed.), *Encyclopedia Britannica* (para. 1). Encyclopedia Britannica, Incorporated. https://www.britannica.com/science/functional-fixedness.

Francis, E. (2014, March 1). Difficulty vs. Complexity: What's the Difference? *Maverik Education.* http://maverikeducation.blogspot.com/2014/03/difficulty-vs-complexity-whats.html

Francis, E. (2018, July 1). What Exactly Is Depth of Knowledge? *Maverik Education.* https://maverikeducation.com/blog/f/what-is-depth-of-knowledge

Hernandez, R., & Pareja, A. S. (2019). *CWRA+ at IMSA - Illinois Mathematics and Science Academy*. Illinois Mathematics and Science Academy. https://www.imsa.edu/wp-content/uploads/2020/09/CWRA-Board-Meeting-18Sept2019.pdf

Hess, K.K (2008, December 17). *Norman Webb's work in test item development*. Center for Assessment. https://www.nciea.org/library/exploring-cognitive-complexity-depth-of-knowledge-part-1/

Hess, K. K., Jones, B. S., Carlock, D., & Walkup, J. R. (2009). C*ognitive rigor: Blending the strengths of Bloom's taxonomy and Webb's depth of knowledge to enhance classroom level processes*. https://files.eric.ed.gov/fulltext/ED517804.pdf

Hess, K. K. (2013). *Cognitive Rigor in Today's Classroom*. Measured Progress Common Core. https://www.asdn.org/wp-content/uploads/CCAP3010_Cognitive_Rigor_in_Todays_Classroom.pdf

Hess, K. K. (2013). *Linking Research with Practice*. Utah Education Network. https://www.uen.org/literacyresources/downloads/linking_research_with_practice.pdf

Hitchcock, D. (Fall 2020 Edition) Critical Thinking. In E. N. Zalta (ed.) *The Stanford Encyclopedia of Philosophy.* The Metaphysics Research Lab Philosophy Department Stanford University. https://plato.stanford.edu/archives/fall2020/entries/critical-thinking

IBM Corporation. (2022). *Global C-suite Series The CEO Study: Own your impact. Practical pathways to transformational sustainability*. https://www.ibm.com/thought-leadership/institute-business-value/en-us/c-suite-study/ceo#

International Institute for Management Development. (2021, June). *IMD World Competitiveness Yearbook* (No. 32). https://www.imd.org/centers/world-competitiveness-center/rankings/world-competitiveness/

Kayembe, C., & Nel, D. (2019). Challenges and Opportunities for Education in the Fourth Industrial Revolution. *African Journal of Public Affairs*, *11*(3), 79--94. https://journals.co.za/doi/pdf/10.10520/EJC-19605d342e

Krathwohl, D. R. (2002). A Revision of Bloom's Taxonomy: An Overview. *Theory Into Practice*, *41*(4), 212–218. https://doi.org/10.1207/s15430421tip4104_2

Lamberti, A. (2020). *An analysis of the complex thinking requirements of the TerraNova and IOWA practice tests in English/language arts for grade 8: A tale of two tests* [Doctoral dissertation, Seton Hall University] Seton Hall University Dissertations and Theses ETDs. https://scholarship.shu.edu/dissertations/2810

Leboff, D. (n.d.). *Student Achievement Definition and Meaning*. Top Hat. https://tophat.com/glossary/s/student-achievement/

Maxwell, J. A. (2013). *Qualitative research design: An interactive approach*. SAGE.

Mayring, P. (2000). Qualitative Content Analysis. *Forum Qualitative Sozialforschung / Forum: Qualitative Social Research*, *1*(2), Article 2. https://doi.org/10.17169/fqs-1.2.1089

McGinnis, D. (2020, October 27). *What is the Fourth Industrial Revolution? The 360 Blog from Salesforce*. https://www.salesforce.com/blog/what-is-the-fourth-industrial-revolution-4ir/

Merriam, S. B. (2009). *Qualitative research: A guide to design and implementation*. San Francisco, CA: Jossey-Bass

Miles, M. B., Huberman, A. M., & Saldaña, J. (2014). *Qualitative data analysis: A methods sourcebook.* SAGE

Moore, T. (2013) Critical thinking: Seven definitions in search of a concept. *Studies in Higher Education*, *38*(4) 506–522. DOI: 10.1080/03075079.2011.586995

Mumford, M. D., Costanza, D. P., Threlfall, K. V., Baughman, W. A., & Reiter-Palmon, R. (1993). Personality variables and problem-construction activities: An exploratory investigation. *Creativity Research Journal*, *6*(4), 365–389. https://doi.org/10.1080/10400419309534493

National Assessment Governing Board. Selected Response. (n.d.). *Selected Response*. National Assessment Governing Board. https://www.nagb.gov/naep-subject-areas/technology-and-engineering-literacy/framework-archive/2014-technology-framework/toc/ch_4/descriptions/descriptions1.html.

New Jersey Department of Education. (n.d.). *Historical Context: Overview of New Jersey's Statewide Testing Program*. New Jersey Department of Education: Assessment: Historical Context: Overview of New Jersey's Statewide Testing Program. https://www.nj.gov/education/assessment/history.shtml

New Jersey Education Association. (n.d.). *High stakes testing*. New Jersey Education Association. https://www.njea.org/high-stakes-testing/

Justia US Law. (n.d.). *2013 New Jersey Revised Statutes: Title 18A - EDUCATION: Section 18A:7C-1 - Commissioner of education to develop a program of standards and guidelines*. Justia Law. https://law.justia.com/codes/new-jersey/2013/title-18a/section-18a-7c-1

Mississippi State University. (2009). *Webb's depth of knowledge guide: Career and technical education definitions.* https://www.aps.edu/sapr/documents/resources/Webbs_DOK_Guide.pdf

Northern Illinois University Center for Innovative Teaching and Learning. (2020). Bloom's taxonomy. In *Instructional guide for university faculty and teaching assistants.* https://www.niu.edu/citl/resources/guides/instructional-guide

Runco, M. A., & Chand, I. (1995). Cognition and creativity. *Educational Psychology Review, 7*(3), 243–267. https://doi.org/10.1007/BF02213373

Schadt, S. (2021). *Critical Thinking and other Higher-Order Thinking Skills.* Center for Excellence in Teaching and Learning. https://cetl.uconn.edu/resources/design-your-course/teaching-and-learning-techniques/critical-thinking-and-other-higher-order-thinking-skills/

Scriven M, Paul R (1987, Summer). *A statement presented at the 8th Annual International Conference on Critical Thinking and Education Reform*. 8th Annual international conference on critical thinking and education reform. http:// www.criticalthinking.org/pages/defining-critical-thinking/766

Scwab, K. (2020, September 22). *The Fourth Industrial Revolution: What it means and how to respond*. World Economic Forum. https://www.weforum.org/agenda/2016/01/the-fourth-industrial-revolution-what-it-means-and-how-to-respond/

Seifert, K. (2009). Educational Psychology (2nd ed.). Global Text. The Saylor Foundation. https://www.saylor.org/site/wp-content/uploads/2012/06/Educational-Psychology.pdf.

Silva, E. (2009). Measuring Skills for 21st-Century Learning. *Phi Delta Kappan*, *90*(9), 630–634. https://doi.org/10.1177/003172170909000905

Solis-Stovall, L. A. (2020). *An Analysis of the Higher Order Thinking Requirements of PARCC Practice Assessments in Grades 3 and 4* [Doctoral dissertation, Seton Hall University] Seton Hall Dissertations and Theses ETDs. https://scholarship.shu.edu/dissertations/2783/

Sydoruk, P. D. (2018). *An analysis of the higher order thinking requirements of a grade 8 online-based English language arts skills program* [Doctoral dissertation, Seton Hall University] Seton Hall Dissertations and Theses ETDs. https://scholarship.shu.edu/dissertations/2495

Thorndike, E. L. (1924). Mental discipline in high school studies. *Journal of Educational Psychology*, *15*(1), 1–22. https://doi.org/10.1037/h0075386

Tienken, C. H. (2020). *Cracking the code of education reform: Creative compliance and ethical leadership*. Corwin.

Tienken, C. H. (2018). Accountability for learning. *Kappa Delta Pi Record*, *54*(2), 56–59. https://doi.org/10.1080/00228958.2018.1443645

Tienken, C. H. (2017). *Defying Standardization: Creating Curriculum for an Uncertain Future*. Rowman & Littlefield.

Tienken, C. H., Colella, A., Angelillo, C., Fox, M., McCahill, K. R., & Wolfe A., (2017) Predicting Middle Level State Standardized Test Results Using Family and Community Demographic Data, RMLE Online, 40:1, 1-13, DOI: 10.1080/19404476.2016.1252304

Tienken, C. H., & Orlich, D. C. (2013). *The school reform landscape: Fraud, myth, and lies*. Rowman et Littlefield Education

Tienken, C. H., Goldberg, S., & Dirocco, D. (2009). Questioning the questions. *Kappa Delta Pi Record*, *46*(1), 39–43. https://doi.org/10.1080/00228958.2009.10516690

Vanderhook, C. A. (2020) *The type of questions being promoted in a 10th grade social sStudies textbook* [Doctoral dissertation, Seton Hall University] Seton Hall University Dissertations and Theses (ETDs). https://scholarship.shu.edu/dissertations/2730

Webb, N. L. (2007). Issues Related to Judging the Alignment of Curriculum Standards and Assessments. *Applied Measurement in Education*, 20(1), 7–25. https://doi.org/10.1080/08957340709336728

Webb, N. L. (1997). *Criteria for alignment of expectations and assessments in mathematics and science education* (Research Monograph No. 6). Washington, DC: Council of Chief State School Officers.

Webb, N.L , Alt, M., Ely, R., & Vesperman, B. (2005). Web Alignment Tool (WAT) training manual. Wisconsin Center for Educational Research, 1(1) 1–177

Whiting, K. (2020, October 21). *What are the top 10 job skills for the future?* World Economic Forum. https://www.weforum.org/agenda/2020/10/top-10-work-skills-of-tomorrow-how-long-it-takes-to-learn-them/

Wiggins, G. (2011). A true test: Toward more authentic and Equitable Assessment. *Phi Delta Kappan*, *92*(7), 81–93. https://doi.org/10.1177/003172171109200721

Wilson, L. O. (n.d.). *Anderson and Krathwohl Bloom's taxonomy revised* - *Quincy College*. Quincy College . https://quincycollege.edu/wp-content/uploads/Anderson-and-Krathwohl_Revised-Blooms-Taxonomy.pdf

Wilson Mulnix, J. (2012) Thinking Critically about Critical Thinking, *Educational Philosophy and Theory*, 44:5, 464-479, https://doi.org/10.1111/j.1469-5812.2010.00673.x

World Economic Forum. (2020, December). *The Global Competitiveness Report Special Edition 2020 How Countries are Performing on the Road to Recovery*. https://www.weforum.org/reports/the-global-competitiveness-report-2020/

Yuan, K., & Le, V.-N. (2014). *Measuring Deeper Learning Through Cognitively Demanding Test Items: Results from the Analysis of Six National and International Exams*. RAND Corporation. https://www.rand.org/pubs/research_reports/RR483.html

Zieky, M., & Perie, M. (2006). *A Primer on Setting Cut Scores on Tests of Educational Achievement*. https://www.ets.org/research/policy_research_reports/publications/publication/2006/dbkw.html

Appendix

Appendix A: Hess' Cognitive Rigor Matrix

Hess' Cognitive Rigor Matrix & Curricular Examples: Applying Webb's Depth-of-Knowledge Levels to Bloom's Cognitive Process Dimensions - *ELA*

Revised Bloom's Taxonomy	**Webb's DOK Level 1 Recall & Reproduction**	**Webb's DOK Level 2 Skills & Concepts**	**Webb's DOK Level 3 Strategic Thinking/ Reasoning**	**Webb's DOK Level 4 Extended Thinking**
Remember Retrieve knowledge from long-term memory, recognize, recall, locate, identify	○ Recall, recognize, or locate basic facts, details, events, or ideas explicit in texts ○ Read words orally in connected text with fluency & accuracy			
Understand Construct meaning, clarify, paraphrase, represent, translate, illustrate, give examples, classify, categorize, summarize, generalize, infer a logical conclusion), predict, compare/contrast, match like ideas, explain, construct models	○ Identify or describe literary elements (characters, setting, sequence, etc.) ○ Select appropriate words when intended meaning/definition is clearly evident ○ Describe/explain who, what, where, when, or how ○ Define/describe facts, details, terms, principles ○ Write simple sentences	○ Specify, explain, show relationships; explain why, cause-effect ○ Give non-examples/examples ○ Summarize results, concepts, ideas ○ Make basic inferences or logical predictions from data or texts ○ Identify main ideas or accurate generalizations of texts ○ Locate information to support explicit-implicit central ideas	○ Explain, generalize, or connect ideas using supporting evidence (quote, example, text reference) ○ Identify/ make inferences about explicit or implicit themes ○ Describe how word choice, point of view, or bias may affect the readers' interpretation of a text ○ Write multi-paragraph composition for specific purpose, focus, voice, tone, & audience	○ Explain how concepts or ideas specifically relate to *other* content domains or concepts ○ Develop generalizations of the results obtained or strategies used and apply them to new problem situations
Apply Carry out or use a procedure in a given situation; carry out (apply to a familiar task), or use (apply) to an unfamiliar task	○ Use language structure (pre/suffix) or word relationships (synonym/antonym) to determine meaning of words ○ Apply rules or resources to edit spelling, grammar, punctuation, conventions, word use ○ Apply basic formats for documenting sources	○ Use context to identify the meaning of words/phrases ○ Obtain and interpret information using text features ○ Develop a text that may be limited to one paragraph ○ Apply simple organizational structures (paragraph, sentence types) in writing	○ Apply a concept in a new context ○ Revise final draft for meaning or progression of ideas ○ Apply internal consistency of text organization and structure to composing a full composition ○ Apply word choice, point of view, style to impact readers' /viewers' interpretation of a text	○ Illustrate how multiple themes (historical, geographic, social) may be interrelated ○ Select or devise an approach among many alternatives to research a novel problem
Analyze Break into constituent parts, determine how parts relate, differentiate between relevant-irrelevant, distinguish, focus, select, organize, outline, find coherence, deconstruct (e.g., for bias or point of view)	○ Identify whether specific information is contained in graphic representations (e.g., map, chart, table, graph, T-chart, diagram) or text features (e.g., headings, subheadings, captions) ○ Decide which text structure is appropriate to audience and purpose	○ Categorize/compare literary elements, terms, facts/details, events ○ Identify use of literary devices ○ Analyze format, organization, & internal text structure (signal words, transitions, semantic cues) of different texts ○ Distinguish: relevant-irrelevant information; fact/opinion ○ Identify characteristic text features; distinguish between texts, genres	○ Analyze information within data sets or texts ○ Analyze interrelationships among concepts, issues, problems ○ Analyze or interpret author's craft (literary devices, viewpoint, or potential bias) to create or critique a text ○ Use reasoning, planning, and evidence to support inferences	○ Analyze multiple sources of evidence, or multiple works by the same author, or across genres, time periods, themes ○ Analyze complex/abstract themes, perspectives, concepts ○ Gather, analyze, and organize multiple information sources ○ Analyze discourse styles
Evaluate Make judgments based on criteria, check, detect inconsistencies or fallacies, judge, critique			○ Cite evidence and develop a logical argument for conjectures ○ Describe, compare, and contrast solution methods ○ Verify reasonableness of results ○ Justify or critique conclusions drawn	○ Evaluate relevancy, accuracy, & completeness of information from multiple sources ○ Apply understanding in a novel way, provide argument or justification for the application
Create Reorganize elements into new patterns/structures, generate, hypothesize, design, plan, produce	Brainstorm ideas, concepts, problems, or perspectives related to a topic or concept	○ Generate conjectures or hypotheses based on observations or prior knowledge and experience	○ Synthesize information within one source or text ○ Develop a complex model for a given situation ○ Develop an alternative solution	○ Synthesize information across multiple sources or texts ○ Articulate a new voice, alternate theme, new knowledge or perspective

Note. Hess's Cognitive Rigor Matrix. From *Cognitive Rigor: Blending the Strengths of Bloom's Taxonomy and Webb's Depth of Knowledge to Enhance Classroom-level Processes* (p. 8) by Hess, et. al, 2009. Publication 2009 by Hess, et al. Reprinted with permission.

Appendix B: Coding Table: Performance Tasks

	A,1	A,2	A,3	A,4	A,6	B,2	B,3	B,4	B,6	C,2	C,3	C,4	C,5	C,6	D,2	D,3	D,4	D,5	D,6
HSA																		X	
HSB																			X
MS																			X

Appendix C: Coding Table: Selected-response questions

	A,1	A,2	A,3	A,4	A,6	B,2	B,3	B,4	B,6	C,2	C,3	C,4	C,5	C,6	D,2	D,3	D,4	D,5	D,6
HS-Q1						X													
HS-Q2						X													
HS-Q3						X													
HS-Q4						X													
MS-Q2						X													
MS-Q3						X													
	A,1	A,2	A,3	A,4	A,6	B,2	B,3	B,4	B,6	C,2	C,3	C,4	C,5	C,6	D,2	D,3	D,4	D,5	D,6
MS-Q4						X													

MS-Q5						X													
MS-Q6						X													
MS-Q7						X													
MS-Q8						X													
MS-Q9						X													
MS-Q10												X							
	A,1	A,2	A,3	A,4	A,6	B,2	B,3	B,4	B,6	C,2	C,3	C,4	C,5	C,6	D,2	D,3	D,4	D,5	D,6
MS-Q11												X							
MS-Q12						X													
MS-Q13						X													
MS-Q14												X							