Critical Thinking: Assessing the Relationship With Academic Achievement and Demographic Factors

A DISSERTATION
SUBMITTED TO THE FACULTY OF THE GRADUATE SCHOOL
OF THE UNIVERSITY OF MINNESOTA
BY

Robyn Wakalua Vierra

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF EDUCATION

Karen Seashore, Ph.D., Adviser
Karen Storm, Ph.D., Adviser

May 2014

ACKNOWLEDGEMENTS

Some say writing a dissertation can be one of the toughest tests in academic life due to the challenges of isolation, procrastination, and perfectionism (Percy, 2014). Luckily, this dissertation journey has not suffered these trials thanks to the many supportive people in my life and role models who have gone before me. Earning her doctorate at 64, Dr. Ruth Borofsky set the bar as a role model for high expectations and achievement. My father, Dr. Rob Borofsky, and sister, Dr. Amelia Borofsky followed in her footsteps, modeling persistence and perseverance. But it would have never been possible if it weren't for the sacrifices of time given to be by Sam, my supportive husband, Nancy, my encouraging mother and babysitter, and the Vierras, my in-laws who spent many family vacations with me posted at the computer. Thanks also to my children Caleb and Tad, who began to think that sitting in front of a Word document amidst a mound of papers was a form a play. The challenge of perfectionism was never an option.

Luckily, I also did not suffer from isolation due to the constant contact and support by my advisers Dr. Seashore and Dr. Storm. Despite being an ocean away, they kept me moving and motivated. Both highly respected in their research fields, they have prioritized being teachers in the truest sense – guiding, supporting, and inspiring their students. I deeply appreciate their comments, questions, and collaboration along the way. Thanks also to my other committee members, Dr. Goh and Dr. Wahlstrom, who showed incredible support and adaptability as I scheduled and rescheduled them from the maternity ward.

Finally, this would not have been possible without the support of Shanghai American School administration and teachers – Mr. Latske, Dr. Mott, Mr. Hendricks, Mr. Musk, and Ms. Wagner – who provided me with access to their students and fielded my barrage of questions.

ABSTRACT

The purpose of this study was to examine the relationship between critical thinking skills and academic performance, and to determine the degree to which demographic characteristics moderate the relationship. The California Critical Thinking Skills Test Middle School Series (CCTST-M series) was administered to assess critical thinking skill levels of students. Academic performance was measured by teacher assigned grades in core subject areas and the Measures of Academic Progress (MAP) test. The demographic factors - gender, tenure at Shanghai American School (SAS), and Culture (native language serving as a proxy for culture data) – were self reported and crosschecked with student records. Data was collected from 297 eighth grade students at Shanghai American School, a high performing American international school located in Shanghai, China.

One-Way ANOVA and Stepwise models were used to examine the relationship between each of the factors and critical thinking. Results showed that grades and MAP test scores were significant predictor variables for critical thinking skills, indicating a strong relationship between critical thinking skills and academic achievement. Gender and tenure at SAS did not yield significant results, and do not moderate the relationship with critical thinking skills. Initial analysis also found culture to be an insignificant variable, except when math performance was factored out, Confucian students scored lower than non-Confucian students in critical thinking. This variance suggests a discipline specificity of critical thinking within some cultures, while also supporting the idea of culturally specific conceptualizations of critical thinking. Additional analysis

also identified a relationship between academic achievement and gender and culture. Females receive higher grades and score higher in the language usage portion of the MAP test. In the mathematics portion of the MAP test, males score higher than females and Confucian students score higher than non-Confucian students.

Results indicate that academic achievement is closely tied with critical thinking and that some variation exists across cultures. Additional research is suggested to further study why these variations, along with differences in academic achievement, exist.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

"It is today we must create the world of the future."
- Eleanor Roosevelt

Statement of problem

Over the last several decades there has been an increasing consensus among policy makers that improving education needs to be a priority. Starting in the early 1980s, various commissions, organized groups and individual commentators have expressed escalating concerns about whether the educational system in the U.S. is prepared to meet future social needs. The initial force behind many of these concerns was the National Commission on Excellence in Education's publication of *A Nation at Risk* (1983). The report stated that: "[T]he educational foundations of our society are presently being eroded by a rising tide of mediocrity that threatens our very future as a Nation and a people" (as cited in Heise, 1994, p. 1). Specifically, it stated "many 17-year-olds do not possess the 'higher-order' intellectual skills we should expect of them. Nearly 40 percent cannot draw inferences from written material; only one-fifth can write a persuasive essay; and only one-third can solve a mathematics problem requiring several steps" (National Commission on Excellence in Education, 1983, p. 5). Thomas Friedman's bestselling book *The World is Flat* (2006), illustrates a continued fear over lagging student achievement. Friedman (2006) asserts that the biggest threat to America is a globalized market which empowers all countries to compete on a level field, thus threatening the stability of American jobs and livelihoods. Friedman (2006) argues for a refocusing of American education around skills, namely creativity, synthesis, analysis,

and holistic thinking, a form of thinking which focuses on noticing patterns and making connections. Additionally, numerous authors have argued that American competitiveness depends on 21st century graduates receiving a skills based education which teaches them to think at higher levels (Browne & Keeley, 1998; Casner-Lotto & Barrington, 2006; P. Facione, 2011; Kuhn, 1999). Education reform is being called to order.

Many of these reforms call for the instruction of 21st century skills to prepare students for a rapidly changing world. 21st century skills include a myriad of skills such as critical thinking, problem solving, collaboration, communication, creativity, and innovation (Greenhill, 2010). This study focuses more narrowly on critical thinking skills, which can be defined as the cognitive skills of interpretation, analysis, evaluation, inference, explanation, and self-regulation (P. Facione, 1990a).

The first wave of change that sought to address the thinking gap, as articulated by *A Nation at Risk* (1983), was President Clinton's Goals 2000: Educate America Act (1994). Goal three of the Educate America Act discusses the need for educators to prepare students for responsible citizenship and productive employment in the modern economy. Specifically, it outlines the goal that:

> "The percentage of all students who demonstrate the ability to reason, solve problems, apply knowledge, and write and communicate effectively will increase substantially; All students will be involved in activities that promote and demonstrate good citizenship, good health, community service, and personal responsibility" (United States Congress, 1994, para. 3).

Implicit in this goal, is a call for an increased "number of college graduates who can think critically, communicate effectively, and solve problems" (D. Resnick & Peterson, 1991, p. 1).

The Goals 2000: Education Act shifted responsibility from states to the federal government and began the process of standards-based reform. States were asked to report on student learning as measured by standards-based tests. Greater emphasis was placed on testing and accountability (Linn, 2000). New standards and assessments were meant to guide teachers in more rigorous instruction (Rothman, Slattery, Vranek, & L. Resnick, 2002).

Unfortunately, there is evidence to suggest that these accountability measures have in fact negatively impacted critical thinking instruction. Standardized tests associated with *No Child Left Behind* tend to exclude higher order thinking skills in favor of simpler cognitive questions (Olson, 2003). The high stakes attached to these tests refocused instruction around low cognitive expectations, and thinking activities have not received priority in classrooms (Conley, 2003; Rothman, et al., 2002). Some critics are hopeful, however, that recent state adoptions of the Common Core State Standards (CCST) will reset a higher bar for rigor and critical thinking (Conley, 2011; McCollister & Sayler, 2010).

Some argue that grading systems have also failed to account for skills learning. Conley (2003) suggests that both state standardized tests and GPAs do not capture the cognitive skills which are truly necessary to be successful later in life. This is evidenced by the fact that GPAs and state standardized test scores are on the rise, while measures of

college success, which include thinking skills, remain steady or have even declined (Woodruff & Ziomek, 2004; Ziomek & Svec, 1995). Both grading systems and state standardized tests may fall short of assessing the skills necessary for the future.

A number of consequences could result from a "non-thinking" curriculum. First, it impacts civic engagement, a core part of democratic society. As Kuhn (1999) explains, critical thinking is an important competency that enables people to "participate as citizens in a democratic society" (p.1). Democracies depend on an educated and critical citizenry to elect candidates, pass propositions, and drive national agendas. To do this, citizens need to know about their choices, know why they are making them, and be able to justify them to others (Ten Dam & Volman, 2004). According to P. Facione (2011), critical thinking also enhances a democracy by empowering citizens to be productive members of society, contributing to, rather than depleting government resources.

Even the Army has taken a serious interest in critical thinking instruction and assessment as a line of defense. The U.S Army Institute (2002) commissioned a report to investigate theories behind critical thinking and the practices that encourage critical thinking in varied contexts, such as battlefields. The report articulates a desire for a clear definition and criteria for critical thinking that could be used to develop training materials and assess success (Cohen, Salas, & Riedel, 2002).

Low thinking skills may also have economic consequences, as governments and business leaders increasingly rely on the thinking skills of citizens as a foundation for a productive economy. Freeley & Steinberg (2008) contend that critical thinking skills are

necessary to succeed in a highly competitive business world. According to Florida (2006) the United States is undergoing a dramatic economic shift away from manufacturing toward a more creative economy, which requires problem solving and independent thinking among other thinking skills. Business leaders must be capable of analyzing market trends, inferring about future demands, and drawing conclusions based on perceived consumer desires (Wagner, 2008). Managing and analyzing a wealth of information and applying it across disciplines are essential skills in the new economy (Wallis & Steptoe, 2006). While the standards-based education movement has shifted education away from a thinking curriculum there is evidence to suggest the economy requires just the opposite.

Despite the arguments for critical thinking as an important life-long skill, there is evidence to suggest its instruction and learning falls short. American students lag behind many other industrialized nations on Program for International Student Assessment (PISA) tests (Programme for International Student Assessment, 2012). The test is particularly challenging to American students because it required them to use critical thinking and problem solving skills in real-world contexts (Kay, 2009). The Washington Post also reported on a national study by the Department of Education, which found American students could read at surface levels, but struggled with deeper thinking and discussion about what they read. 80% of third graders, over half the seventh graders, and 36% of 11th graders did not meet expectations for critical thinking on a critical thinking reading assessments (Vobejda, 1988). Conley (2007) suggests that in order to close the domestic as well as the global achievement gap students must have key cognitive skills, a

category under which critical thinking falls, when they enter college. Unfortunately, he argues, students are falling short. This is evidenced by Ali & Jenkins' (2002) study, which found that 46% of all University of California students required remedial coursework upon entry (as cited in Conley, 2007). In response, all University of California students must now take a critical thinking course as part of their mandatory requirements. Additionally, Kwon (2008) found that undergraduate students suffered from high levels of anxiety when confronted with vast amounts information. They struggled to identify and evaluate relevant information from a numerous sources, which is one of the skills required of the ideal critical thinker.

Even upon graduation, university and college students still do not meet future employers' thinking expectations. The *Workforce Readiness Report Card* (2006), which surveyed more than 400 Fortune 500 companies, found that the new workforce was under prepared for the new economy and lacking in basic knowledge and applied skills. Among those applied skills necessary, critical thinking and problem-solving were at the top of the list (McLester & McIntire, 2006). Additionally, a study of newly registered nurses found they did not meet job expectations for entry-level critical judgment (Del Bueno, 2005). Not only are students not learning to think, but the California Commission on Teacher Credentialing reports that teachers don't know how to teach them to do so (Paul, Elder, & Bartell, 1997).

In order to better prepare students for college, the workforce, and effective citizenship, the literature suggests a growing need to rethink the measures used to define high school and middle school success. Though some see promise in newer Common

Core aligned tests (Conley, 2011), many state standardized tests currently focus on lower level knowledge and do not prioritize or assess the skills necessary to be successful in the 21st century. According to the Partnership for 21st Century Skills and the American Association of Colleges of Teacher Education (2010) students need a "mastery of the 21st century skills such as critical thinking, problem-solving, communication, and collaboration and creativity and innovation" (p. 6). Unfortunately, the movement for accountability and testing has not propelled this agenda forward.

Defining Critical Thinking

Critical thinking has strong philosophical roots, tracing back to the works of Socrates, Plato, and Aristotle. Socrates focused on the idea of self-examination, arguing for reflection, analysis, and openness to criticism as a necessity for human fulfillment. Plato and Aristotle believed in the use of logic and questioning to make reasoned judgments that would ultimately lead to personal freedom. More recently, Ennis (1985) describes critical thinking as "reflective and reasonable thinking that is focused on deciding what to believe or do" (p. 45), McPeck (1981) explains it as "the propensity and skill to engage in an activity with reflective skepticism" (p. 8), and Paul (1992) defines critical thinking as "the intellectually disciplined process of actively and skillfully conceptualizing, applying, analyzing, synthesizing, and/or evaluating information gathered from, or generated by, observation, experience, reflection, reasoning, or communication, as a guide to belief and action" (p. 22). Finally, philosopher and teacher, Harvey Siegel (1980) approaches critical thinking as the "ability and the

willingness to be objective, impartial and non-arbitrary, based on evidence" (p. 4). Philosophers tend to look at the hypothetical critical thinker's characteristics, emphasizing qualities or standards of thought (Lai, 2011).

Educators have also sought to use their expertise to examine critical thinking. Though Dewey was a philosopher and psychologist, he was also very interested in educational reform. His conception of critical thinking, or reflective thinking as he called it, was aimed at developing thinking skills among students. He focused on the educator's role in teaching students to question and seek evidence to guide them in favor or against what they previously believed to be true (Dewey, 1910). Many educators also associate Bloom and his taxonomy of learning with critical thinking. Teachers are often encouraged to use the taxonomy to bring students to higher levels of learning by asking them to analyze, synthesize, and evaluate information (Bloom, 1956). The educational approach to critical thinking often focuses on the role of the teacher for critical thinking instruction.

Most recently cognitive psychologists have weighed in on the critical thinking debate using current brain research on learning. Psychological approaches tend to view critical thinking more as a scientific process, focusing on individual actions or behaviors rather than the sum of the whole (Bailin, 2002; Gelder, 2005). Willingham (2008) defines critical thinking as "seeing both sides of an issue, being open to new evidence that disconfirms your ideas, reasoning dispassionately, demanding that claims be backed by evidence, deducting and inferring conclusions from available facts, solving problems,

and so forth" (p. 8). Critical thinking in the cognitive sciences tends to focus around meta-knowing, a larger umbrella of thinking.

There are clear areas of disagreement in the debate to define critical thinking, but general skills pertaining to critical thinking, those of reasoning, problem solving, and judging, are common throughout. There is also basic agreement on the necessity of critical thinking dispositions, the habits of mind or inclination for critical thinking, in order to exemplify the ideal critical thinker, though they may not be included in a definition for critical thinking skills. Finally, most scholars across the disciplines would suggest that some basic knowledge is a prerequisite for critical thinking. Critical thinkers need something to think critically about.

The extent of discipline specificity and transferability of critical thinking is, however, at the heart of the disagreement. While there is general agreement that some knowledge is important, scholars range from suggesting basic knowledge (P. Facione, 1990d) to requiring extensive expert knowledge for deeper thinking (National Research Council, 2004). The role of knowledge, then, impacts the domain specificity and transferability of critical thinking skills across contexts. If expert knowledge is necessary, then one can only be an excellent critical thinker in some subjects. If, however, just a basic knowledge is necessary, critical thinking skills can transcend across disciplines and be used to gain further knowledge elsewhere. The role of knowledge and the discipline specificity of critical thinking will impact assessment design and future instruction.

The lack of consensus over a definition for critical thinking has impeded progress toward a thinking curriculum and its assessment (Brunt, 2005; Lai, 2011; Silva, 2009). In 1990 the American Philosophical Association attended to this issue by convening a panel of experts from various fields to discuss a common definition for critical thinking. The resulting Delphi Study defined critical thinking as ""purposeful, self-regulatory judgment which results in interpretation, analysis, evaluation, and inference, as well as explanation of the evidential, conceptual, methodological, criteriological, or conceptual considerations upon which that judgment is based" (P. Facione, 1990 p.3). This definition separates skills and dispositions while also recognizing their interdependence. It also recognizes the necessity of knowledge for critical thinking, but not to the extent that critical thinking cannot transfer across disciplines. This definition presupposes that individuals may be more experienced critical thinkers in some areas than others, but there is universality to the ideal critical thinker.

Statement of Study Purpose

The purpose of this study is to examine the relationship between critical thinking skills and academic performance of middle school students at an American international school in China, and to determine the degree to which student demographic characteristics moderate the relationship.

Research Questions

1) What is the level of critical thinking, as measured by the California Critical Thinking Skills Test, of eighth grade students at an American international school in Shanghai, China?

2) To what extent is there a relationship between critical thinking and student achievement, as measured by grades and MAP test scores?

3) Do student levels of critical thinking vary in terms of the following demographic variables?

- Gender
- Tenure at Shanghai American School
- Native Language

Significance of Study and Research Implications

This research can inform education boards, administrators, accreditation agencies or any stakeholders who work to improve schools. The results of this study specifically provide feedback on Shanghai American School's fulfillment of its mission and vision, while also informing other schools on how to address critical thinking in their grading and testing practices. Schools looking to find data on student critical thinking may choose to adopt critical thinking assessments, general standardized tests like the MAP or simply analyze their grades and grading practices. The results of this study can recommend or contest the reliability of using existing assessment methods to provided information on student critical thinking skills.

This study may also encourage educational leaders to adopt more formal critical thinking assessments as they work to develop vision statements, weigh initiatives, and delegate money. School leaders may decide to use the information from critical thinking tests or evaluations to set goals and guide school agendas. Critical thinking data may, for example, suggest areas for teacher professional development, teacher evaluation, and curriculum review. Leaders may also use data to identify educators who demonstrate excellence in critical thinking instruction and capitalize on their expertise and knowledge in organizing learning communities. This study provides an example of critical thinking assessment at one school, but has practical and policy implication for other schools wishing to support critical thinking instruction and learning.

The information gained in this study also adds to the literature on the relationship between gender and critical thinking. Critical thinking has been accused of male bias (Thayer-Bacon, 1993), yet the data on gender differences is mixed. Some studies find women outperform men on critical thinking tests (Srinivasan & Cooks, 2005), others find men outperform women (Leach, 2011; Simon, 1974), and still others have found no gender correlation at all (Ben-Chaim, Ron, & Zoller, 2000; El Hassan & Madhum, 2007; P. Facione, 1990c; P. Facione, Sanchez, N. Facione, & Gainen, 1995). This study attempts to address the gaps in research on gender and critical thinking, particularly in secondary school.

Results also provide information on the relationship between culture and critical thinking. Critical thinking as part of the educational experience has taken on a new dimension with the globalization of education. Not only are businesses moving

overseas, but transnational education is also on the rise. Students are studying abroad in record numbers and distance-learning programs can be found in abundance (Sun & Boncella, 2007; Turner, 2006). American educational values and approaches to learning are being exported to international schools around the world, while at the same time foreign students are applying to American academic institutions in record numbers (Helms, 2008; "More Korean children sent to study abroad alone," 2009). There is a greater need for cross-cultural understanding of thinking and learning styles.

Native language in this study serves as a proxy for culture, and languages are grouped into two broad categories: Confucian (for Cantonese, Mandarin, Korean, Japanese, and Taiwanese speakers) and Anglo (for English speakers). When it comes to critical thinking, Asian students are often criticized for lacking problem-solving and reasoning abilities (Turner, 2006). Results contribute to the literature on culture and thinking styles. White a number of studies have addressed the relationship between critical thinking dispositions and culture, few studies have focused on the skills aspect of critical thinking. This study helps address this research gap. Results also advise approaches for teachers to improve their instruction in culturally diverse classrooms. A relationship between critical thinking skills and culture, for example, indicates that teachers need to provide explicit instruction on critical thinking skills expectations and study methods (Turner, 2006). No correlation, however, suggests that dispositional differences may be to blame for stereotypes of the Asian student. In such a case, teachers may then model and instruct on how to formulate questions, challenge authority, and contribute to discussions (Ku & Ho, 2010). A defined relationship between culture and

critical thinking skills will guide instruction and curriculum decisions particularly in diverse international school settings.

Second, the research offers a better understanding of the relationship between critical thinking skills, dispositions, and culture. Asian students in foreign contexts can often be observed sitting quietly while diligently scribing lecture notes dictated by the professor. These students are sometimes criticized for their lack of participation and passivity (Biggs, 1996). It is unclear, however, whether these criticisms are based on stylistic judgments of Asian thinking and learning dispositions or if students actually lack critical thinking skills. A number of studies have identified a negative relationship between critical thinking dispositions and students of Confucian heritage (Ip, et al., 2000; Ku & Ho, 2010; Tiwari, Avery, & Lai, 2003). Little has been done, however, to assess cultural differences in skills, so it is unclear whether dispositions are being negatively affected by lack of skills, skills are not learned due to lack of disposition toward critical thinking, or whether the two are mutually exclusive. Results aid this discussion and perhaps lead to less intellectual imperialism.

Third, the analysis of native language, years of schooling at SAS, and critical thinking skills data informs the debate on the malleability of thinking in diverse populations. Hofstede, Hofstede, & Minkov (2010) argue that cultural values and orientations to learning are learned early in life and change slowly and across generations. Other research, however, has found that Chinese students adapt quite well to new learning expectations and environments. Little research has been done in the way of looking at learning and thinking of Asian students studying in American schools within

Asia. Do newly admitted students exhibit lower critical thinking skills than students who have attended SAS for longer periods of time? Does time in an American school correlate with critical thinking skills or as Hofstede et al. (2010) assert, are views of authority, uncertainty avoidance, and approach to learning longstanding and largely influenced by family and culture? Results of this study do not provide formative data on this relationship but will further the debate.

Context of Study

This study was conducted at Shanghai American School (SAS), an independent, non-profit school sponsored by the U.S. Consulate in Shanghai. It is the only American school in Shanghai, but one of numerous international schools. SAS is the largest expatriate school in Shanghai, and the largest in China. It is the second largest international school in the world. Between the Puxi and Pudong middle school campuses, there are 1,051 students. The majority of SAS teachers are American (57%), with Canadian teachers being the second most represented nationality (15%), and Chinese the third (9%). The student body represents 40 different nationalities based of origin of passport.

SAS offers a rigorous academic American curriculum, as well as numerous extracurricular activities. SAS is an authorized Advanced Placement and International Baccalaureate school. An average of 98% of the students go on to university, with the majority matriculating to universities in the USA. At the middle school level all students are enrolled in language arts, social studies, math, physical education, a foreign language

course, and music. There are also numerous enrichment activities such as interscholastic sports programs, China Alive, and abundant service learning opportunities. The National Middle School Association principles guide the Middle School program, “focusing on providing developmentally appropriate challenges in cognitive, social, emotional, physical and moral development” ("Shanghai American School middle school program," 2012, para. 1). There are no courses specific to critical thinking learning, but there is an understanding that is threaded throughout the entire curriculum.

SAS is primarily funded by student tuition. Tuition costs vary by grade level, but the average cost is approximately $22,000 per student. At one time tuition was paid primarily by business employers, but in recent years more parents are paying out of pocket. The operating budget for 2012-2013 was roughly $64,000,000 and is funded almost entirely by student tuition.

In March of 2011 the Board of Directors formally approved a new mission and vision statement put forward by a panel of teachers, parents, students and administrators. The SAS mission states that Shanghai American School “inspires in all students a lifelong passion for learning, a commitment to act with integrity and compassion, and the courage to live their dreams” (Shanghai American School, 2012, para. 3). The SAS core values assert a belief that:

- Embracing diversity enriches individuals and communities
- Acts of compassion and generosity of spirit create a better world
- When individuals take responsibility for their own decisions, they are empowered to make positive impact

- Each individual has intrinsic value and the potential to contribute to society
- Collaboration is key to overcoming complex challenges and achieving common goals
- Integrity is the foundation of enduring relationships, quality institutions and well-functioning communities
- As global citizens we have a duty to care for the earth and its inhabitants to ensure the well-being of humankind
- Creativity, critical thinking and a lifelong passion for learning are essential to personal fulfillment and to meet the challenges of the future.

  (Shanghai American School, 2012, para. 2).

The new mission and strategic objectives reflect a philosophical shift in the organization. While the previous focus was on becoming "a leading school in Asia," it is now about empowering students to act with character and pursuit of their personal passions.

Theoretical Construct

This study is based on the conceptual understand of critical thinking as defined by the Delphi Report. The Delphi research project, sponsored by the American Philosophical Association, was a two-year study, which included 46 experts with backgrounds in critical thinking education, research and/or assessment. After five rounds of facilitated Delphi method discussions, the group was able to define the characteristics of critical thinking skills, dispositions, and the ideal critical thinker. Their consensus understanding of the skills and dispositions for critical thinking can be found in the key

terms. This construct of critical thinking is non-domain specific and integrates content knowledge as a means to as well as an end of critical thinking.

Key Terms

Critical thinking. Critical thinking is the purposeful, self-regulatory judgment which results in interpretation, analysis, evaluation, and inference, as well as explanation of the evidential, conceptual, methodological, criteriological, or contextual considerations upon which that judgment is based (P. Facione, 1990a, p. 3).

Critical thinking skills. Critical thinking skills are the cognitive skills associated with critical thinking. Interpretation, analysis, evaluation, inference, explanation, and self-regulation are the six cognitive skills central to critical thinking (P. Facione, 1990a).

Critical thinking disposition. Critical thinking disposition is the approach to critical thinking and life in general. It is the probing inquisitiveness, a keenness of mind, a zealous dedication to reason, and a hunger or eagerness for reliable information. Systematic, inquisitive, judicious, truth seeking, confidence of reasoning open-mindedness, and analytical behavior are the subsets associated with a critical thinking disposition (P. Facione, 2011, p.10).

California Critical Thinking Skills Test (CCTST). The California Critical Thinking Skills Test is a family of instruments used to collect data on critical thinking skills. The CCTST M-Series, designed specifically for students in grades 3-9, measure 6 dimensions associated with critical thinking skills: analysis and interpretation, inference, evaluation and explanation, deductive reasoning, inductive reasoning, and numeracy.

The test is based on the Delphi consensus definition of critical thinking published by the APA (P. Facione, Gittens, N. Facione, & Winterhalter, 2011).

Grades. Grades are reported as a percentage out of one hundred. Grades are assigned using teacher judgment and school academic standards. Grades at Shanghai American School should reflect only the skill level of a student, not his or her attitude or learning dispositions. Non-skill items are reported on the Learner Profile and are not included in the academic final grade.

Measures of Academic Progress (MAP) Test. The Measures of Academic Progress Test is an instrument used to measure student's academic level. It measures the areas of reading, mathematics, and language usage ("MAP overview," 2012).

Assessment. Assessment is the process of estimating a student's progress toward an objective and using that information to help students continue their learning (*This we believe: Successful schools for young adolescents*, 2003).

Evaluation. Evaluation is the process of using data and standards to judge the quality of progress or level of achievement (*This we believe: Successful schools for young adolescents*, 2003).

Native language. Native language is defined as the primary language spoken at home.

Delimitations and Limitations

This study is an investigation of the relationship between critical thinking skills and other forms of assessment such as the MAP test and teacher assigned grades. Since

both the MAP test and grades are meant to assess student skill levels, critical thinking skills rather than dispositions are being measured. While dispositions are an essential part of being a critical thinker, they are not included in the scope of this study. Nonetheless, dispositions are shown to correlate strongly with thinking skills and an examination of student disposition levels would be an excellent follow-up to this study.

Another delimiter of this study is the focus on assessment, not instructional practices or pedagogical approaches to support critical thinking. Assessment can be used to inform areas for instruction or evaluate the effectiveness of particular instructional practices, but this inquiry comes after an initial assessment. Critical thinking assessment is a necessary first step before instructional analysis. Therefore, best practices in critical thinking instruction are not a focus of discussion.

This study has limitations of both design and methodology. First, the sample includes only one grade-level, at one school, at a single point in time (N=282). A wider set of subjects across time would increase the validity and reliability of the study. Secondly, Critical thinking skills rely heavily on cognitive development (Kuhn, 1999; National Research Council, 2004). An argument can be made that an attainable score on a critical thinking test is limited by a relatively young sample of students. While the sample group may not be at the peak of their critical thinking potential, they are at an age were it is being developed. Middle school students are ripe for learning 21st century skills because they have the core knowledge and skills on which to build (Kay, 2009). Additionally, since test results are being compared and normed to students of similar ages, the results will still highlight exemplary or deficient critical thinking skills for the

specified age group. Conclusions regarding middle school assessment and demographic differences will be valid.

There are also some limitations in regards to the instrument. Multiple-choice tests have been criticized for being inauthentic, not motivating student performance, and failing to address the thinking processes behind a marked answer. Though multiple-choice tests are considered an indirect assessment they do have wider acceptance of validity and reliability in the academic community than other forms of assessment (Erwin, 2000). These issues are not specific to critical thinking tests, rather they are generic to multiple choice testing in general. Regardless, multiple-choice tests are widely used and accepted testing instruments. Additionally, the CCTST was designed for and validated in American schools. While SAS is an American school it does have a much larger Asian student population than most American schools in the United States. Differences in background beliefs between the test makers and the test takers may affect results. Background or cultural differences could cause justifiably different answers than those identified as correct (Ennis, 1993). Multiple-choice tests have been criticized for bias and favoritism of students from specific backgrounds. As of now, however, there is no research to support the existence of cultural bias on critical thinking assessments (Erwin, 2000).

Another limitation of this study is the use of native language as a proxy for culture. Student passport, mother's passport, and father's passport are also considered in the analysis. It is difficult to identify a single or predominant culture in this sample group. Culture can be defined in a variety of ways and it is particularly difficult for third culture

kids or kids of mixed heritage. Many students hold American passports and have spent some time in the United States, but appear of Chinese origin and speak Mandarin at home. There are other students who have grown up in numerous international locations and have taken on attributes of various cultures during their formative years. To avoid the ambiguity of a defined culture, native language serves in its place. However, the grouping of individuals into a singular cultural category can be problematic, along with terms used for the categorizing. Many of the students come from mixed backgrounds, are bilingual, and even struggle to define a single culture for themselves. Thus, requiring students to name a native language to represent their cultures is a necessary flaw in the research.

Summary

Excellent education is an essential good for an individual as well as society. Learning to think critically is one of the many necessary educational learning outcomes that schools must foster and develop to promote an enlightened, productive and successful citizenry. Today, more than ever before, educators are being asked to provide evidence of these desired learning outcomes to stakeholders who want to know whether their investments are valuable. Assessment has come to forefront of educational discussions and initiatives. There are numerous ways to assess student learning, but those focusing on skills are increasingly being valued over those that measure content learning. Many schools have implemented new testing programs and grade realignment

initiatives with the specific goal of skills assessment. The extent to which critical thinking skills are included in these assessments, however, is unclear. This study seeks to address this research gap by investigating the relationship between current assessment practices and critical thinking performance. School leaders and educators need to understand the extent to which students are learning to be critical thinkers. The data on critical thinking may already be embedded in current systems or it may be an area for further inquiry. Regardless, obtaining this data is an important first step to making informed instructional, program, and budgetary decisions.

While direct assessment is the first step to identifying areas for critical learning, it is also important to understand the factors beyond just instruction, which may impact critical thinking and learning. Some of these factors may be gender, the curriculum, and culture. Understanding the weight of these factors is of particular importance in an increasingly global and multicultural world. This study investigates the relationships between critical thinking skills and gender, tenure at SAS, and culture. Understanding the impacts of each of these factors helps educators improve their programs and instruction.

Overview of Chapters

In Chapter 2, a review of the literature introduces the necessity of critical thinking as a 21st Century skill, while also putting it in a historical context for how philosophers, educators, and psychologists have defined or addressed it over time. This review then goes on to discuss why and how critical thinking should be assessed. Finally, a focused

literature review of the research questions is presented. Subsequently Chapter 3 discusses the quantitative methods and methodology used in this study. A discussion of the sample, procedures, and context of the study, as well as a review of the different data sets collected and their respective reliability and validities are included. Subsequently, Chapter 4 presents the findings from the quantitative analysis. Here, descriptive data is presented first, providing context and rationale for how the subsequent one-way ANOVA and Stepwise regression models are analyzed. Critical thinking is used primarily as the dependent variable in all models, but additional analysis of academic achievement is also offered. Chapter 5 summarizes the findings presented in Chapter 4, offering interpretations and implications for the results. Finally, limitations of the study and suggested areas for future research are recommended.

CHAPTER 2: LITERATURE REVIEW

" Fix reason firmly in her seat, and call to her tribunal every fact, every opinion. Question with boldness even the existence of a god; because, if there be one, he must more approve of the homage of reason, than that of blindfolded fear."
-Thomas Jefferson

Introduction

This chapter begins by making a case for the importance of critical thinking as an educational goal, drawing on literature from Greek philosophers and reaching into the present, the argument for critical thinking rests primarily on Western values of personal freedom and democratic ideals, although it also has roots in Eastern philosophy.[1] The context of this study, an American international school with graduates primarily going on to study at American institutions, is a rationale for focusing on Western theory and research.

The conceptualization and articulation of critical thinking has evolved and continues to evolve - with new common understandings and emerging brain research on learning. Ultimately, this study will describe critical thinking as the "purposeful, self-regulatory judgment which results in interpretation, analysis, evaluation, and inference, as well as explanation of the evidential, conceptual, methodological, criteriological, or conceptual considerations upon which that judgment is based" (P. Facione, 1990, p.3). This chapter will first examine *what* is critical thinking and the debate over how to define critical thinking. Next, it will look at the *how* of critical assessment – its rationale and methods – rather than on the large body of research that investigates instruction and

[1] Critical thinking is valued around the world, though not necessarily in the same sense or for similar purposes.

pedagogy that augment critical thinking in students. Attention will also be paid to the relationship between critical thinking, grades, and standardized tests in the literature. Finally, this review discusses demographic variables, gender, and ethnicity, as they relate to critical thinking. This is the area least researched and results of this study could make a positive contribution in this area particularly. By better understanding student levels of critical thinking and how they relate to student background factors, educators and policy makers can begin to make more educated decisions for improved critical thinking education and assessment.

Why Critical Thinking Matters

Today we are inundated with more information than ever before in human history. News podcasts, social networking sites, and television advertisements are just a few of the many sources one might encounter in a single hour. It is with this information that many of us create our personal truths and understandings of the world. The good critical thinker can navigate the tidal wave of content – analyzing, weighing, and sometimes discarding information. Poor critical thinkers, however, are bound by preconceived thoughts, the beliefs of their social networks, and those of the media. Critical thinking is asked of us on a daily, even minute-by-minute basis. Simply navigating a grocery store and the labels of "natural goodness" or "whole grain" requires critical skills. One must ask how whole grain is defined or what constitutes an "all natural" label. Even valid news sources can lead the uncritical astray. Imagine picking up the morning paper, for example, and seeing a recent article on the high incidence of children diagnosed with

ADHD in the Palo Alto area. The unreflective or uncritical Palo Alto resident may instantly panic in fear that her children are now at risk for such a diagnosis. While this is a possibility, there are also many other explanations that must be weighed. It requires one to ask questions, analyze, and evaluate. How is ADHD being defined by this study? Might there be a relationship between the skills necessary for computer engineering, the main economy of Palo Alto, and ADHD? Perhaps Palo Alto simply has a reputation for excellent services, thus drawing parents of ADHD children to the area. Without critical thinking we cannot dig into the realm of information with confidence and power, nor are we able to make rational, thoughtful choices.

Asch's conformity experiment (1950) and Milgram's prisoner experiment (1961) are two examples that illustrate the importance of critical thinking. Asch's (1950) experiment looked at how conformity makes people do things that conflict with what they know to be true. Asch asked subjects to look at an image and answer a series of obvious questions, but only after posed confederates answered the question first. When Asch's confederates answered incorrectly, the experimental subject was also more likely to answer with an incorrect answer. The results revealed that individuals are less likely to exhibit critical thinking when in opposition to a group. Participants of the experiment feared being humiliated or ridiculed for presenting an answer that was different from those of the confederates.

Similarly, Milgram's (1961) research required subjects to administer an electric shock to the learner in the next room. The subject could hear the screams and shouts of the learner but was told to continue regardless. In many of the trials, the subject

complied with the orders and administered lethal level shocks to the learner. This study demonstrated people's willingness to obey authority figures and institutions. The overwhelming number of participants who administered lethal shocks to the Milgram's confederate puzzled and horrified the public. Many of the participants were educated middle-class citizens, yet most revealed a lack of acute critical thinking and conviction. These two studies not only illustrate the power of conformity and obedience, but they also highlight the need for individuals to be able to independently analyze and evaluate a situation. Individuals who acquire critical thinking skills at an early age and practice using those skills throughout their lifetimes may be less likely to conform to group think or to obey immoral orders of perceived authority figures.

Arguments for critical thinking are rooted in Western philosophy. Famous philosophers Socrates, Plato, and Aristotle all embraced the human quest for truth and freedom through critical thought. Socrates is famous for having said, "the unexamined life is not worth living" (Plato, 1892/1928, p. 41). He argued that reflection, analysis, and openness to criticism are necessary for human fulfillment. Life, otherwise, is superficial and unworthy according to Socrates (O'Reilly, 2010). Socrates is also famous for his method of using questioning, inquiry, and debate to deepen student thinking about a topic. This is now known as the Socratic Method or the Socratic Seminar, both of which are commonplace is education today. In a Socratic Seminar students are encouraged to use dialogue to identify fallacies, break down what they already know, and come to new understandings. The teacher is a mentor, pointing out flaws in argument, but does not define the discussion. This method is embraced for its encouragement of

skepticism and critical thought, both of which Socrates exemplified. In the end, Socrates most likely died for questioning socially accepted beliefs on power and politics.

Both Plato, a student of Socrates, and Aristotle, a student of Plato, expanded on and formalized Socrates' methods, using logic as the basis for philosophical argument. The idea that logic is necessary to making reasoned judgments became the foundation of Western philosophy. Plato and Aristotle both argued for the importance of reasoning and critical thought in obtaining personal freedom. In Plato's *Allegory of the Cave* (1928), he wrote that cavemen "have their legs and necks chained so that they cannot move, and can only see before them being prevented by the chains from turning around their heads" (p. 514). He argued that cavemen are tied to their traditions and the knowledge handed to them. They are not free or thoughtful thinkers. Plato believed that we should not be tied to traditions or upbringings, that we must instead create individual realities by questioning and creating personal interpretations. Similarly, Aristotle has been quoted as saying that "life according to reason is best and pleasantest, since reason more than anything else is man. This life therefore is also the happiest" (as cited in Irwin, 1999, p.38). The human's job is to reason just like an eye's job is to see; it is through the execution of this role that one finds fulfillment. Critical thinking and an argument for its supreme value can be traced back Greek philosophy.

Western philosophers are not, however, the only ones to stress the importance of a reasoned and critical life. Confucius also argued for the importance of reflective thinking, as he called it, which is at the heart of the Confucian tradition. Learning, according to Confucius, is not simply the process of memorizing and gaining a basic

understanding, it is "studying extensively, enquiring carefully, pondering thoroughly, sifting clearly, and practicing earnestly…" (as cited in Lee, 1996, p. 35). His conceptualization of learning emphasizes inquiry and open-mindedness as requirements to being a reflective thinker. Learning in this sense is necessary to the cultivation of the self as well as the success of society.

John Dewey, the most influential American philosopher of education, believed in the centrality of critical or reflective thinking to the educational experience and its role in establishing personal freedom. Aloofness or indifference, Dewey argued, creates dependence, which should be avoided. Beliefs and judgments are often results of unconscious learning from those around us and we need to learn to sort and create independent thoughts not based on authority. Dewey argued that educational importance should not be placed in *what* students learn but *how* they learn – through discussion, questioning, and analysis (Dewey, 1966).

Critical thought is not only a part of the human quest for truth and fulfillment, but it is a means to freedom. Educational philosopher Paulo Freire is likely the most prominent advocate for the freedom argument. In *Pedagogy of the Oppressed*, he argues that students are not deposits for information, but rather creators of information who must develop a "critical consciousness" (Freire, 1974). Every citizen has the right and responsibility to go on his or her own quest for knowledge. Education, according to Freire, is about creating your own knowledge, and this knowledge will make you free from the oppressor.

The contemporary educational philosopher Robert Siegel also believes in the liberating forces of critical thinking: "The self-sufficient person is a liberated person… free from the unwarranted and undesirable control of unjustified beliefs" (Siegel, 1988, p. 58) . Critical thinking is a source of empowerment that frees the individual from the confines of collective thought.

The argument for critical thinking continues into modern practical society. The discussion of critical thinking heated up in the 1960s, which is not incidentally the same time Americans were concerned about communism and the safety of democracy. Americans began thinking about how to protect the citizenry from propaganda and the threat of totalitarian rule. More than ever before, education for critical thinking was given greater attention (Thayer-Bacon, 2000). The Berlin Wall has since fallen and the U.S. maintains peaceful relations with Russia, but the necessity of critical thinking and its role in education are still central to preserving democratic institutions. As Purdy (1992) argues, it is a moral imperative to teach critical thinking to safeguard from indoctrination and judgment of claims.

Successful democracies depend on the abilities of citizens to use critical thinking skills (Bailin & Siegel, 2003; Kuhn, 1999; Weinstein, 1991). According to Gutmann (1995), the goal of a liberal education is educate students not only for citizenship but also for independent thinking. In order to address social, environmental, and political issues of the future, thinking skills must improve along with the growing complexities of the problems (Richmond, 1993). Citizens are regularly asked to interpret documents, analyze arguments, evaluate policies, infer about the repercussions of laws, explain their views,

and regulate themselves with consciousness. Opportunities exist for citizens to critically participate in their democracy through both formal and informal means. For example, jurors are asked to critically judge arguments made by lawyers, weigh evidence, and ultimately decide on the fate of the accused. Voters are asked to analyze social and economic issues and then vote on a candidate to represent their opinions. It is up to citizens to decide whether a railroad should be built, homosexuals should get married, or nature reserves should be preserved. Citizens are given significant rights with the assumption they will use them responsibly and critically. The Bill of Rights gives citizens the right to protest, to carry a gun, and to protect their privacy in the absence of a warrant. Democracies become legitimate when citizens can articulate differing points of view and also question the views or positions presented to them (Gutmann & Thompson, 2004). These responsibilities and rights exemplify the American democracy but cannot be dealt with casually and without deep reflection and critical thought.

It is through education that citizens learn how to thoughtfully participate in a democracy. Ten Dam and Monique (2009) argue that critical thinking is crucial to competent participation in a plural and democratic society. According to The Carnegie Forum, schools must provide:

> "A deeper understanding necessary for self governing society… It must enable the citizens of this country to make informed judgments about the complex issues and events that characterize life in advanced economies at the end of the 20th century. The cost of not doing so may well be the gradual erosions of our democratic birthright… The focus of schooling must shift from teaching to

learning, from the passive acquisition of facts and routines to the active application of ideas to problems" (as cited in Walsh & Paul, 1986, p. 7).

Schools should breed effective democracy, but as some argue, this is not always the case. In 1983, "A Nation at Risk" reported on the high risk of deficient thinking skills. The report warned that the lack of thinking skills will result in greater disfranchisement from material rewards and full participation in national life (Milton & Harvey, 1983). In 1984, the Paideia Proposal reported that the lack of discussion and opportunities for higher order thinking in classrooms were undemocratic (Adler, 1984). Excellent instruction in school is necessary for the protection of democratic ideals.

A critical and well-educated society is not only good for democracy but also good for the whole; P. Facione (2011) argues, "becoming educated and practicing good judgment …is better than enduring the consequences of making bad decisions and better than burdening friends, family and all the rest of us with the unwanted and avoidable consequences of those poor choices" (p. 2). An educated and participatory citizenry is less likely to engage in crime, fall into poverty, and suffer from health issues. According to Ten Dam and Monique (2004), critical thinking enables citizens to contribute to society. Critical thinkers are more likely to be successful academically and later in life (Jenkins, 1998; R. Williams, Oliver, Allin, Winn, & Booher, 2003). Such success relieves dependence on the social net.

Employers are also demanding more thoughtful, prepared graduates who can think critically. A survey of accounting faculty and practitioners found that they rate critical thinking among the top three most important skills for graduates (Sharifi,

McCombs, Fraser, & McCabe, 2009). Business recruiters said critical thinking skills are among the most important basic skills they look for in prospective employees (Hopkins, Raymond, & Carlson, 2011). Nursing programs recognize the need for employees to be capable of making sound professional judgments and are investing millions in educating and assessing critical thinking skills (Chenoweth, 1998; Girot, 1995). Wagner (2010) reported that less than 25 percent of surveyed employers thought college educated employees had excellent knowledge or skills, and that approximately 50 percent believed their employees hired right from high school were deficient in their preparation for basic jobs. Job requirements will change significantly by the time students are ready to fill them, but students must be educated to fill those jobs now.

The only way to address this issue is to educate a force of thinkers. Individuals who can question and think critically will always be able to adapt and adjust to new demands in the job market (Wagner, 2008). Business leaders have called on education programs to focus less on technical skills and more on critical thinking skills that will equip employees to deal with the challenges of the 21st century (Jenkins, 1998). In a survey of business leaders from various economic sectors, Wagner (2008) found that business leaders identified critical thinking and problem solving as the most important skills for employees. In the 21st century job market, they explained, employees need more than a knowledge of the field – they also need an ability to manage a wealth of information and to ask the right questions to solve new problems (Wagner, 2010). The need for workers to analyze and think critically is not new, but it is more important in today's rapidly changing market (Huitt, 1995; Silva, 2009).

Critical thinking not only prepares students for the workforce, but it also sets them up for future academic success. Numerous studies have shown that critical thinking predicts future academic success (D. Allen & Bond, 2001; Jenkins, 1998; Scott & Markert, 1994; Williams, 2003; R. Williams, et al., 2003). If students are expected to be critical thinkers in college and beyond, they must be given opportunities to hone their skills as early as kindergarten (DeVoogd, 2006). If one falls behind in critical thinking skills, not only are academic courses more difficult, but improving critical thinking skills becomes even harder (Williams & Stockdale, 2003). It is a slippery slope.

By the time students get to high school, students are already expected to have some level of critical thinking skills. College preparatory courses, like the International Baccalaureate (IB) and Advanced Placement (AP), use critical thinking in their tests (Rothschild, 1999; Taylor & Porath, 2006). In 2007, 1.4 million students took an AP exam (College Board, 2008) and 71,130 students took the IB exam (International Baccalaureate North America, 2007). Student enrollments in these college preparatory courses can predict college acceptance (NCEA, 2010; Santoli, 2002) and increase chances of being admitted to a selective universities or colleges (Western Interstate Commission for Higher Education, 2006).

Success in college and university is also tied to critical thinking skills. Conley (2007) found that success in post-secondary school requires a combination of key cognitive skills, academic knowledge and skills, and academic behaviors. These key cognitive skills are similar to critical thinking in that they emphasize analysis, interpretation, precision and accuracy, problem solving, and reasoning. The Association

of American Colleges and Universities expects students to have the "abilities to analyze, communicate, and integrate ideas; and effectiveness in dealing with values, relating to divers individuals, and developing as individuals" (as cited in Gaff, 2004, p. 4). Bok (2008) reports that 90% of graduate level instructors rank critical thinking as one of the most important things students can learn in undergraduate education. If critical thinking predicts academic success and is tied to performance in AP and IB courses, which in turn correlates with college acceptance, then critical thinking skills are essential and must be developed early on.

Critical thinking instruction is clearly important. It is a valuable part of the human experience, offering freedom of thought and personal growth. It is a habit of mind that supports life long excellence. Critical thinking is also of institutional value; it protects a healthy democracy, promotes economic prosperity, depletes dependence on government resources, and is necessary to a productive workforce. Most educators would nod in agreement, recognizing its immense value. While many assume it is integrated and learned within the core curriculum, the extent to which this is true is unclear (Halpern, 2001). Critical thinking matters and must be treated so.

What Is Critical Thinking[2]?

If educators, policy makers, job creators, and academics all agree critical thinking is not only valuable but imperative, then the first step is to define what is actually meant by critical thinking. For critical thinking to be part of the educational experience and

[2] The discussion of how to define critical thinking focuses on the philosophical and educational research, touching briefly on the cognitive psychological approach.

instructional standards, educators must have clarity over what they are expected to teach. Without a common definition, educators are fumbling, teaching what they *think* constitutes critical thinking; meanwhile, students are confused by the various meanings and expectations put forth by different teachers. Educators not only define critical thinking skills differently, but evidence shows they also define it inaccurately (Beyer, 1984). Educators should not be the brunt of the blame, however. Textbook makers, curriculum design specialists, and educational philosophers are all known to ascribe to their own unique definitions of critical thinking. How critical thinking is integrated (or not integrated) into the curriculum, pedagogical approaches, and the weight it is given in the educational experience depend on a clear definition. While the critical thinking debate is lively and interesting, progress is impeded by a lack of agreement (Brunt, 2005; Lai, 2011; Silva, 2009).

Not only is it difficult to teach critical thinking without a clear definition, but it can't be assessed. Stakeholders will only know if critical thinking instruction has been successful and where to focus continued efforts if they can reliably assess student critical thinking abilities. In *Understanding by Design (2005)*, Wiggins and McTighe argue for backwards design, which is the process of using assessments to design curriculum, performance assessments, and instruction. Assessment drives practice. Such assessment requires a clear definition and criteria of what exactly is meant by critical thinking. Critical thinking instruction requires assessment, but assessment requires a clearly articulated definition with criteria.

In an effort to define critical thinking, the works of educational philosopher John Dewey are a good place to start. He is regarded by some as the father of critical thinking, though he referred to it as "reflective thinking" (Fisher, 2001; Kuhn, 1999; Nelsen & Seaman, 2011; Sawaya, 2012). In his book *How We Think* (1910), Dewey discusses the importance of the "scientific attitude of mind" and "reflective thinking." While Dewey uses the term reflective thinking rather than critical thinking, he does contrast being reflective with being uncritical. Dewey's sense of reflective thinking is similar, then, to what we now call critical thinking. Dewey defines reflective thinking as "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, p. 9). The reflective thinker does not always turn to what he knows, beliefs or habits he has picked up consciously or unconsciously, but rather suspends judgment in order to search new materials and evidence before accepting a conclusion. Dewey emphasizes the need to carefully arrive at decisions or opinions, to provide reasons for these opinions, and to thoroughly understand the implications of them (Fisher, 2001). Dewey's take on reflective thinking has poignant implications for today's schools and the conceptualization of critical thinking.

In educational context, it is common to hear Benjamin Bloom's work references when talking about critical thinking (Kennedy, Fisher, & Ennis, 1991). According to Bloom, higher order thinking requires analysis, synthesis, and evaluation, with comprehension and application filling in the bottom rung of lower order thinking skills (Bloom, 1956). Bloom's influential taxonomy of learning objectives is commonplace in

educational discussions and teacher trainings and is widely cited in educational literature. Though the higher taxonomy levels of analysis, synthesis, and evaluation, are often used to define critical thinking skills, scholars in the critical thinking field argue that this is a misrepresentation of critical thinking (Ennis, 1993; Huitt, 1998; Krathwohl, 2002; Paul, 1985a). While the taxonomy is a useful tool for developing assessments and guiding instruction, it should not represent the order in which critical thinking skills are acquired or how they interact with one another (Ennis, 1993). According to Bloom's hierarchy of learning, once an individual gets to the top levels of thinking, he or she can be thought of as a critical thinker. This, however, is not how most scholars view critical thinking (Ennis, 1980; P. Facione, 1990d). Most critical thinking experts would argue that critical thinking skills are inter-related, each "level" relying on another (Ennis, 1993; P. Facione, 1990d). Critical thinking is not a neat set of steps with a hierarchical order of acquisition. Rather, the skills interact with one another in a more fluid and complex manner (Ennis, 1993; Richard Paul, 1985a). Paul (1985a) also argues that knowledge is misrepresented in the taxonomy as a straightforward and simple process of memorization. Acquiring knowledge, he says, involves complex processes of sorting, questioning, and connecting information. Recent studies on the brain and learning support Paul's criticism, concluding that knowledge is constructed and retained with the use of complex thinking skills (National Research Council, 2004; L. Resnick, 2010; White & Frederiksen, 1998). More recently Bloom's Taxonomy has been revised to reflect new discoveries in brain research, with knowledge being changed to remembering, comprehension to understanding, and evaluation to creating, as well as allowing for greater flexibility

between the levels (Forehand, 2010). These changes address some of the discussed critiques but not necessarily to the extent that it should be used to define critical thinking skills (Krathwohl, 2002; Sternberg, 1986).

Robert Ennis, who has devoted himself to critical thinking since the early 1950s, was one of the first contemporary philosophers to construct a concrete definition for critical thinking. Critical thinking, according to Ennis (1980), "is reasonable reflective thinking focused on deciding what to believe or do" (p. 180). He also provides details for his definition, outlining what a critical person should be able to do:

1) Judge the credibility of sources.

2) Identify conclusions, reasons, and assumptions.

3) Judge the credibility of an argument, including the acceptability of its reasons, assumptions, and evidence.

4) Develop and defend a position on an issue.

5) Ask appropriate clarifying questions

6) Plan experiments and judge experimental designs.

7) Define terms in a way appropriate for the content.

8) Be open-minded.

9) Try to be well informed.

10) Draw conclusions when warranted, but with caution (Ennis, 1993, p. 180)

For years this definition has served as the backbone of critical thinking. The problem, according to Paul (1992) and Siegel (1988), is that the definition and set of requirements focus on what a person should be able to do but doesn't necessitate a willingness to

perform such activities. Ennis has often been criticized for his failure to include the "critical spirit," "dispositions," or the "character traits" that involve the application of critical thinking skills in his definition. Critics argue that the application of skills is a requirement to being a critical thinker, and thus Ennis's definition is too narrow (Paul, 1992; Siegel, 1988).

These character traits and skills are combined in what educational philosopher John McPeck (1981) calls "reflective skepticism." Critical thinking, he argues, is "the propensity and skill to engage in an activity with reflective skepticism" (p. 8). More central to McPeck's sense of critical thinking, however, is his emphasis on discipline specificity. He argues that "thinking is always thinking about X, and that X can never be 'everything in general' but must always be something in particular. Thus the claim "I teach my students to think is at worst false and at best misleading" (McPeck, 1981, p. 4). According to McPeck, it is impossible to have excellent general critical thinking skills; one has to think critically about topics in which he or she is highly knowledgeable. Though critics might argue that he is right at a fundamental level, the extent of his argument leaves no room for general skills or holistic ability (P. Facione, 1990d). According to McPeck's sense of critical thinking, one can be a phenomenal critical thinker in biology but mediocre at best when it comes to politics. Critics argue that his dependence on knowledge significantly limits the breadth of critical thinking.

Fellow philosopher Richard Paul is highly critical of McPeck's work, arguing that critical thinking must be more constructive, open minded, and multi-textured than McPeck affords (Paul, 1985b). Instead he proposes a "strong sense" definition of critical

thinking to address the problems he sees with both McPeck's and Ennis's work. Paul distinguishes between critical thinking in the "weak sense" and the "strong sense." Critical thinking in the "weak sense," he argues, is simply having the skills and being able to use them when asked, but a "strong sense" critical thinker not only has the skills but uses them in daily life. Thus, critical thinking is "the intellectually disciplined process of actively and skillfully conceptualizing, applying, analyzing, synthesizing, and/or evaluating information gathered from, or generated by, observation, experience, reflection, reasoning, or communication, as a guide to belief and action" (Paul, 1992, p. 22). Paul believes critical thinkers can ask questions, summarize an opposing argument, debate an issue, and not rely on others to create personal truths. No thought, he argues, should be permanent, and beliefs should constantly evolve (Paul & Binker, 1990). This definition is also not without its criticisms. It has been criticized for having too much reliance on character or dispositions (Ennis, 1993; Norris, 1985), for lacking detail (Ennis, 1993), and having relativistic tendencies (Siegel, 1988).

Philosopher and teacher Harvey Siegel offers yet another opinion. Siegel (1980) takes a philosophical approach in defining critical thinking as the "ability and the willingness to be objective, impartial and non-arbitrary, based on evidence" (p. 4). According to Siegel critical thinking is an educational ideal to which teachers have an ethical and moral responsibility. Without critical thinking, he argues, one is bound by control of others. Unlike McPeck, he says critical thinking is implicit in education because understanding requires analysis and evaluation. To understand a math theorem to be true, one must use inference and evidence. To study science, one must use

judgment against the evidence. Mastery of a subject matter cannot be done without critical thinking. Siegel criticizes Paul for his "relativistic tendencies," McPeck for his discipline-specific approach, and Ennis for his failure to include action or willingness in his definition. Siegel's approach is moved by reason and is more abstract in that it lacks concrete criteria for critical thinking. He recognizes this deficiency, but calls on others to address this issue. Siegel argues that the ponderance of critical thinking is not the job of the educator or the philosopher alone. In order to develop the concept of the ideal critical thinker, other domains must weigh in. He asks what does the ideal critical thinker look like in politics, management, and economics? The critical thinking debate, he argues, should engage a wider community of scholars and professionals.

The cognitive psychological approach to critical thinking. The discussion of critical thinking thus far draws primarily on philosophical and educational constructs of critical thinking. Greater empirical research on learning and cognition over the last decade has prompted cognitive psychologists to weigh in on the discussion (National Research Council, 2004). Instead of focusing on critical thinking, cognitive psychologist tend to focus on metacognition or meta-knowing. Though these terms are not synonymous with critical thinking, there are significant overlaps. Metacognition is often described as "thinking about thinking" (Blakey & Spence, 1990, para. 1) or "knowing about ones knowing" (Kuhn, 1999, p. 179). According to Kuhn (1999), metacognitive, metastrategic, and epistemological skills are the core of critical thinking. Metacognitive is about looking at where information comes from and its intended audience (Bransford, et al., 2005). Metastrategic is about using strategies to judge and apply knowledge and

skills to new information and situations (Kuhn, 1999). It is putting information in a broader context, which allows one to accept it as true or reject it. Epistemological understandings can be best connected to critical thinking dispositions. They refer to the belief that knowing is worth thinking about. Together, these meta-knowing skills put individuals in control of their knowing, allowing them to decide what they believe and why – a theme which is also central to critical thinking (Kuhn, 1999).

Cognitive psychologist have also defined critical thinking as "seeing both sides of an issue, being open to new evidence that disconfirms your ideas, reasoning dispassionately, demanding that claims be backed by evidence, deducting and inferring conclusions from available facts, solving problems, and so forth" (Willingham, 2008, p. 8). Halpern (1998) characterizes it as "the use of those cognitive skills or strategies that increase the probability of a desirable outcome" (p. 450). Psychological approaches tend to focus on the individual actions or behaviors of critical thinkers rather the sum of the whole (Bailin, 2002; Gelder, 2005). Experts from the philosophical camp have critiqued these definitions as being too procedural and based only on what can be easily observed (Bailin, 2002; Gelder, 2005; Sternberg, 1986).

The role of knowledge is also critical to the psychological perspective. Meta-knowing focuses on the control and accumulation of knowledge for learning. With knowledge, students can start problem solving at higher levels (Berliner, 2001). Thus, knowledge or knowing about a topic becomes a prerequisite to thinking critically about that topic. There is a growing consensus among psychologists as well as educators that to be an expert thinker, under which critical thinking would fall, there is a need for

subject matter knowledge (Bransford, et al., 2005; Fisher, 2001; Kennedy, et al., 1991). Brain research shows that thinking abilities are developed through content learning in regular curriculum, but discussion, questioning, paraphrasing, and interpreting are the activities to develop long-term learning. Thinking skills are developed by thinking about a topic in which they have expert knowledge (Bransford, et al., 2005).

Knowledge can enhance thinking, but thinking can also lead to greater knowledge. A report by the National Research Council (2004) concluded that while knowledge is an important factor in learning, it is not useful unless it is used as a foundation for deeper understanding. Students learn not by being told new information, but by building on, connecting to, or questioning what they already know. If a student must apply information and draw relationships with other information, he or she is more likely to acquire new knowledge in a related area. Expert thinking about a topic leads to more knowledge in a related field (National Research Council, 2004). More contentious, however, is whether these thinking skills can lead to greater knowledge and thinking across disciplines. Kuhn (1999) argues that critical thinking can be used to acquire new knowledge in areas of less expertise by drawing conclusions or applying similar protocols for accepting or rejecting knowledge. Adey, Csapo, Demetriou, Hautamaki, & Shayer (2007) agree, arguing for a notion of general cognitive ability that can be applied to a variety of contexts. According to Healy (1990) the act of making connections between subjects is what promotes more meaningful critical thinking. Psychologists, however, do not generally accept this argument.

Arriving at a consensus definition. While this discussion of "what is critical thinking?" has not exhausted all the approaches to critical thinking, it does represent major differences and opinions in the field. Only recently have experts from cognitive sciences, socio-cultural research, and educational philosophy come together for cross-disciplinary conversations on the topic (Nelsen & Seaman, 2011). Differences over a definition are obvious. One commonality, however, is clearly present. Nearly everyone believes critical thinking should be thought of as both skills and dispositions. While Ennis does not explicitly discuss this in his original definition, he does go on to discuss the importance of action in becoming the ideal critical thinker in later works (Ennis, 1993). The first part of being a critical thinker – having the skills – refers to ability. It is the cognitive aptitude for critical thinking. The second part – dispositions – refers to the inclination or desire to apply these skills. This is often called the "critical or reflective spirit" or, in psychology, the "epistemological understandings." Critical thinking skills are useless, many argue, if one does not use them. Some scholars separate the skills and dispositions as two separate definitions that would then define the ideal critical thinker. Others believe they are too interdependent to stand-alone. Regardless, they are both part of the conceptualization of the ideal critical thinker (P. Facione, 1990d).

In response to the growing awareness about critical thinking, coupled with ambiguity over how it was being defined, a group of forty-six international experts was convened in 1990 to discuss the issue in an effort to develop a consensus definition. Though all the participants were considered experts in the field of critical thinking, they drew from different backgrounds, such as philosophy, education, the social sciences, and

the physical sciences. Major names in critical thinking, such as Ennis, Brady, Norris, Parry, Paul, and Stiggins, were among the experts on the panel.

After five rounds of facilitated discussions and debate, the American Philosophical Association (APA) published the Delphi Report. The consensus definition is as follows: "We understand critical thinking to be purposeful, self-regulatory judgment which results in interpretation, analysis, evaluation, and inference, as well as the explanation of the evidential, conceptual, methodological, criteriological, or contextual considerations upon which that judgment was based" (P. Facione, 1990d, p. 3). Critical thinking, they determined, consists of six cognitive skills with corresponding sub-skills.

The first skill, I*nterpretation*, is explained by the ability to categorize, decode significance, and clarify meaning. In order to understand these skills, it might be easier to put them in a real-life context, one that everyone is prone to experience. Let us take the scenario of the supermarket. While shopping, a critical thinker might exhibit his ability to interpret by describing the clientele of the store based on the types of products being sold, the pricing, and the proximity to other stores. He might also recognize a problem with the traffic patterns of the store and describe them without bias. Through interpretation, one can "comprehend and express the meaning or significance of a wide variety of experiences, situations, data, events, judgments, conventions, beliefs, rules, procedures, or criteria" (P. Facione, 1990d, p. 13).

The next skill is *analysis* and can be explained as the ability to examine ideas, identify arguments, and break down arguments. While still shopping, the analytical thinker may ponder whether to buy the organic or non-organic rice. He thinks of what he

has read about organic labeling requirements, the dirty dozen, pesticide use, and the fees producers pay to obtain such labels and begins to form a decision or conclusion regarding his choice. He might recognize that organic is a healthy choice, while weighing it against the knowledge of the organic labeling business and the fact that rice production uses few pesticides to begin with. He weighs arguments, identifies assumptions, and draws conclusions with evidence.

The third skill, *evaluation*, is defined as the ability to "assess the credibility of statements or other representations which are accounts or descriptions of a person's perception, experience, situation, judgment, belief, or option; and to assess the logical strength of the actual or intended inferential relationships among statements, descriptions, questions or other forms of representation" (P. Facione, 1990d, p. 15). The shopper might judge the credibility of the sign outside claiming that this particular store has the best prices in town. He might also judge the contradiction between his cereals "100% whole grain" label and the fact that just one gram of fiber is listed under the nutrition facts. Evaluation requires one to assess claims and arguments.

The fourth critical thinking skill, *inference*, involves querying evidence, conjecturing alternatives, and drawing conclusions. While on the way to the store, the shopper may have noticed a wholesale market opening just down the road from his local grocery store. He might draw on what he knows about business and marketing to make predictions about how his local store will adapt to stay afloat with this new competition. He will consider that perhaps his store might offer specials, focus on providing specialty items, or devise a point system for frequent shoppers. Perhaps he might even develop a

plan to gather more information that can inform the store's strategy. Inference requires one to pull from different areas of knowledge and generate thoughtful hypotheses and conclusions.

Explanation, as defined by the Delphi experts, is "to state and to justify that reasoning in terms of the evidential, conceptual, methodological, criteriological, and contextual considerations upon which one's results were based; and to present one's reasoning in the form of cogent arguments" (P. Facione, 1990d, p. 18). At the checkout stand the cashier tells the shopper that the other brand of pineapples are on sale and a few dollars could be saved if he gets the other brand. The shopper has already considered the fact that he wants to eat the pineapple as soon as he gets home, requiring a fully ripe fruit, which this one is. He recognizes he has not selected the cheaper or sweeter variety, but he believes the locally sourced but more expensive pineapple he chose might prove more delicious while also allowing him to support local producers. The cashier may not want to hear this explanation, nor do the customers lining up, but the point is that our critical shopper can thoughtfully explain his reasoning and cite evidence.

Finally, *self-regulation* allows the critical thinker to hone his critical thinking skills with self-examination and self-correction. The shopper has reached the end of his critical shopping experience and looks over his purchases. He knows he has bought the smaller milk container, which is impractical for his large family and he reflects on why he made this silly decision. He notices the packaging, particularly that the colors and glass bottle remind him of his childhood fridge and the delicious experience of sitting down to his mother's cooking with a cold glass of milk. Additionally, he takes in the

happy-looking cows on the label that help him reconcile his discontent with the milk industry and the treatment of cows. The shopper reflects on his purchase and evaluates his reasoning, coming to the conclusion that next time he will go with the other, larger and more practical brand.

This definition of critical thinking attends to a wide range of audiences, across disciplines, yet gives concise clarity to the term and is currently being used to inform assessments, program standards, graduation requirements, policy initiatives and more (N. Facione & P. Facione, 1997; P. Facione, 1990a; P. Facione, et al., 2011). It is important to note the skills are not listed hierarchically and are not necessarily acquired in a linear fashion. In addition, it incorporates the underlying assumption that reasoning is a complex process that can be reactive as well as reflective (P. Facione & N. Facione, 2007). An excellent critical thinker need not exhibit all the skills at one time, but should be versatile in them all. Though content knowledge is not part of critical thinking, the experts agree some content knowledge is important to applying critical thinking skills (P. Facione, 1990d; Hofer & Pintrich, 1997; Sternberg, 1986).

Critical thinking dispositions versus skills. The ideal critical thinker must not only have thinking skills, but also the dispositions, or inclination, for critical thinking. The Delphi committee could not reach a consensus on whether dispositions should be part of the definition. One-third of the participants argued "critical thinking refers only to the cognitive skills and dispositions, but not affective dispositions" (P. Facione, 1990d, p. 21). The majority, on the other hand, argued that "affective dispositions constitute part of the meaning of CT… and that these dispositions flow from, and are implied by the very

concept of CT, much as the cognitive dispositions are" (P. Facione, 1990d, p. 21). Therefore, the definitions were separated in order to distinguish between the two.

Critical thinking dispositions are divided into *approaches to life and living in general* and *approaches to specific issues, questions or problems*. The critical thinking dispositions are further divided into seven dimensions: truth seeking, open-mindedness, analyticity, critical thinking, self-confidence, inquisitiveness, and maturity (P. Facione, 1990d). The disposed critical thinker leads himself as well as others to critical thinking. Dispositions are often modeled rather than explicitly taught and can be more difficult to assess than skills (Strahan, 1986; Tishman, Jay, & Perkins, 1993). However, dispositions are an important element in discussing the ideal critical thinker, they are not the focus of this study.

Together the critical thinking skills and dispositions come together to describe the ideal critical thinker. The panel defined the ideal critical thinker as follows:

> "Habitually inquisitive, well-informed, trustful of reason, open-minded, flexible, fair-minded in evaluation, honest in facing personal biases, prudent in making judgments, willing to reconsider, clear about issues, orderly in complex matters, diligent in seeking relevant information, reasonable in the selection of criteria, focused in inquiry, and persistent in seeking results which are as precise as the subject and the circumstances of inquiry permit" (P. Facione, 1990d, p. 3).

An individual need not exhibit every subset of critical thinking skills and dispositions at one time, but must be versatile in employing them as required in different contexts. Critical thinking requires maturity and purposeful development. It is developed with

time, but one must develop critical thinking skills and dispositions early in order to reach this critical thinking ideal (J. Brown, 1983; P. Facione, 1990d).

Defining critical thinking for this study. There are many ways to describe or define critical thinking. Conceptually, critical thinking is related to other thinking, such as reflective judgment, problem framing, higher-order thinking, logical reasoning, decision-making, problem solving, the scientific method (Giancarlo & P. Facione, 2001), creative thinking, motivation, and metacognition (Cohen, et al., 2002; Lai, 2011). While all these types of thinking are interrelated and overlap significantly (Kuhn, 2000; Lai, 2011), critical thinking can be defined as an independent dimension of thinking (P. Facione, 1990d). For the purpose of this study, the definition brought forth by the American Philosophical Association's Delphi Report will be used. Not only was it developed with the input of scholars from a range of fields including the sciences, philosophy, and education (P. Facione, 1990d), but the clear criteria suit the educational assessment purposes of this study (Cohen, et al., 2002). The categories and subcategories in the Delphi definition are valuable to educators as they seek to design curriculum and assessment. Additionally, the definition allows for the transferability of critical thinking across disciplines and is not domain specific. While knowledge is needed to employ critical thinking skills, this definition does not presuppose expertise in a specific area. Since this study seeks to measure overall critical thinking skills across disciplines, the APA definition is appropriate.

Critical Thinking Assessment: Why?

The movement for assessment and accountability has been central to education in the 21st Century. It is no longer enough to assume students are learning and making progress toward curriculum standards. Voters, policy makers, administrators, parents and teachers are demanding more evidence to support claims of learning. Educators and policy makers are being asked the fundamental question of what should students be learning and how do we know they are doing so?

Current standardized assessments have, however, been criticized for their narrow focus on content knowledge and for testing just the easiest standards. Olson (2003) found that items requiring harder cognitive processes are being omitted in favor of lower and simpler cognitive questions on state tests. Despite the rhetoric of teaching 21st century skills, accountability measures are not matching the educational ideals being put forth (L. & Zurawsky, 2005). Low-level testing gives teachers little incentive to teach higher cognitive thinking (Diamond, 2007; L. Resnick, 2010). As a result, instruction is refocused on lower cognitive tasks (Rothman, et al., 2002), and curriculum is being reshaped to meet test expectations (Conley, 2003). The disconnect between standards and assessment has teachers and students forgoing higher level thinking in favor of the simplest standards.

More recently there has been significant progress with regards to testing and the inclusion of more cognitively demanding activities on standardized tests. The Common Core State Standards, which nearly all states have now adopted, are more rigorous standards for thinking and learning. Common Core Standards integrate Webb's Depth of

Knowledge and Bloom's Taxonomy to raise the bar, focusing less on knowledge and recall, and more on creation and extended thinking (Hess, Carlock, Jones, & Walkup, 2009). Assessment of the Common Core is still evolving and states will likely be adopting the new tests in coming years. The College-readiness Performance Assessment System is another assessment example that measure cognitive skills essential to college success. The test is based on cognitive theory that says learning occurs when students "construct new knowledge based on what they already know and believe" (Conley, Lombardi, Seburn, & McGaughy, 2009, p. 6). The educational climate has permanently changed to embrace standards-based assessments, and improving the tests and standards to include a wider array of desired educational outcomes is necessary to accommodate the policy realities outlined in Chapter One.

Assessing critical thinking is not only important because it sets instructional priorities, but because the data can provide valuable feedback to stakeholders on student learning. Though international schools are not subject to the same testing requirements as public schools in the United States, assessing student learning and growth is still good educational practice (Chappuis & Stiggins, 2002). Critical thinking data can be used by teachers to improve instruction, by administrators to make informed program decisions, and by parents and voters to assess the return on their investments. Accreditation agencies can also use it to hold schools accountable to their mission, vision, and curriculum standards (Erwin, 2000). Another use for critical thinking data is to improve the reliability and validity of grading as a means of reporting student learning. Research on grading shows a disconnect between reports of student learning and actual standards-

based achievement (Mansfield, 2001; Milton & Harvey, 1983). Critical thinking assessment can help teachers and administrators assess and reflect upon their grading practices. Finally, assessment data can be used as another data point to assess the validity of other instruments, while also providing additional information to inform the discussion of student learning.

Most educators would adamantly agree critical thinking is an essential part of the educational experience and an important life-long skill. Yet critical thinking is often not pursued with the same persistence nor given as much weight as lower-level cognitive tasks (L. Resnick & Zurawsky, 2005). Critical thinking assessments will help teachers, administrators and stakeholders make informed decisions for student learning that are aligned with instructional standards and expectations.

Critical Thinking Assessment: How?

Assessment and accountability proponents have sometimes excluded thinking assessments with the argument that thinking is too abstract to measure (Ennis, 1993). This is not the case. There are a growing number of assessments that can measure both knowledge and skills, recognizing the interconnected nature and necessity of these two learning goals (National Research Council, 2004). Numerous assessments for critical thinking skills and dispositions are available, but only a few have been highlighted for discussion as they pertain to the focus of this study. This study is based on a theoretical construct that students should be able to think across disciplines, and outside of academic contexts and apply critical thinking to their everyday lives. Thus, the following

instruments and assessment strategies are all non-subject specific, measure all aspects of critical thinking, are available for middle and/or high school students, and are well known in the field.

Watson-Glaser Critical Thinking Assessment (1925). The Watson-Glaser Critical Thinking assessment, designed by Watson and Glaser, was one of the first critical thinking tests developed. The first version was originally published in the 1930s and has served as a benchmark for subsequent tests. This test consists of forty multiple-choice questions to assesses the following subsets: inference, recognition of assumptions, deductions, interpretation, evaluation of arguments. The test is written at a ninth grade reading level but can be used with adults as well.

The Cornell Critical Thinking Test (1985). Ennis, Millman, and Tomko designed the Cornell Critical Thinking Test. The test is based on the definition that critical thinking is "the process of reasonably deciding what to believe or do" (Ennis, 1993, p. 180). Level X of the test measures induction, deduction, credibility, and identification of assumptions with a series of seventy-one multiple-choice questions. It can be used for grades four through twelve, and a Level Z test can be given to college-level students.

Ennis-Wier Critical Thinking Essay Test (1985). Ennis and Wier developed a test for grades seven and up to assess critical thinking through written argument. Students are asked to formulate and articulate a convincing argument based on a scenario. The advantage of an essay test is that it allows for numerous correct answers and can represent a student's ability to think through an argument. The fact that it relies entirely

on written ability, however, can be problematic for poor writers. The test is also time consuming to score and more susceptible to assessor biases.

California Critical Thinking Skills Tests (1992). Designed by P. Facione, the CCTST family of tests measures skills associated with critical thinking. The skills test assesses analysis and interpretation, inference, evaluation and explanation, deductive reasoning, inductive reasoning, and most recently, numeracy. These scales are based on the Delphi Expert Panel's critical thinking definition discussed earlier. The original CCTST is for college level students, but a newer M-series is now available for younger students. Separate tests for critical thinking dispositions are also available.

Portfolios. Portfolios are collections of student work used to represent learning and growth. The portfolio process generally involves student reflection and discussions on their learning. Portfolios can be beneficial because the reflection process allows teachers to assess critical thinking as well as instruct at the same time (McMullan, et al., 2003). The downsides are that they are often done wrong, are time consuming, and may not include all types of learning outcomes (N. Facione & P. Facione, 1996).

Performance evaluations. Performance evaluations are an authentic way to incorporate critical thinking assessment into the classroom (Wiggins, 1989). Like portfolios, they also offer opportunities for learning and instruction (Silva, 2009). Critics argue that performance evaluations often don't include all elements of critical thinking and are not valid assessments because they are usually rehearsed and presented in groups (N. Facione & P. Facione, 1996) .

Each of these methods and instruments has its flaws. Multiple-choice tests cannot distinguish between a guess and a systematic analysis. Students may be rewarded with a correct answer even if they did not employ critical thinking skills. At the same time, multiple choice tests reward only one correct answer. There is no reward for a student who uses critical thinking to arrive at an alternative answer (Norris, 1988). With further probing, this student may be able to thoughtfully explain a rationale for why he or she selected an incorrect answer, but this is impossible with a multiple-choice format.

A more holistic approach to critical thinking assessment might include performance evaluations, portfolios, essays, or anecdotal notes from classroom observations. It can be argued that these forms of assessment are more authentic and capture students in their natural environments. Additionally, students are likely to exhibit higher motivation in authentic contexts where performance is tied to a grade or something real (Tombari & Borich, 1999; Wiggins, 1989). These measures, however, are far more susceptible to bias and do not share the same levels of reliability and validity as a tested instrument which use a multiple-choice format (N. Facione & P. Facione, 1996). Performance tasks generally allow students time to practice, and teams of students can work together to prepare. It is difficult to identify what a student can do independently (Ennis, 1993). The main problem with essays is that they require considerable time to grade and even more time if graders need training to align their grading practices (Erwin, 2000). The International Baccalaureate program uses essay tests, but multiple readers score them in order to increase validity. This can be costly and time consuming. Secondly, essay tests may not necessarily assess all the subsets of

critical thinking. It is difficult to address each critical thinking skill within an essay format and even more difficult to weigh the subsets equally (Erwin, 2000).

Critical Thinking and Academic Achievement

Numerous studies, generally at the college level, highlight a strong relationship between critical thinking and achievement as measured by GPA (Cabrera, 1992; Garett & Wulf, 1978; Steward & Al-Abdulla, 1989; R. Williams, et al., 2003; R. Williams & Stockdale, 2003). R. Williams and Stockdale (2003) found high critical thinkers generally received As and Bs and very rarely received Ds or Fs. Low critical thinkers, however, averaged Bs and Cs. Steward and Al-Abdulla (1989) observed a high correlation between overall critical thinking skills and GPA, with a particularly strong relationship in the interpretation subset. In a study of over 1,100 college students, scores on the CCTST critical thinking skills test significantly correlated with college GPA (P. Facione, 1990b). Similar studies by Vendrely's (2007), Jenkins (1998) and P. Facione & N. Facione (1997) found similar results. This is not surprising given most scholars, particularly McPeck, argue that some content knowledge is needed for critical thinking. The more knowledge one has, the higher the level of critical thinking that can occur (Brunt, 2005; N. Facione & P. Facione, 1996). The research does not clearly indicate, however, whether critical thinking leads to higher academic success or higher academic success leads to higher critical thinking. It is most likely a combination of the two (Williams & Stockdale, 2003).

Standardized tests also correlate with critical thinking skills. Williams and Stockdale (2003) reported a strong relationship between scores on the Watson-Glaser

Critical Thinking Appraisal, the CCTST, and the American College Testing (ACT) scores. The CCTST technical report indicated a strong correlation between critical thinking and both the math and verbal sections of the Scholastic Assessment Test (SAT) (P. Facione, 1990a). Verbal abilities particularly influence critical thinking performance. Clifford, Boufal, and Kurtz (2004) identified a relationship between the *Wechsler Adult Intelligence Scale*, the V*erbal Comprehension Index, and critical thinking skills. Unfortunately, the bulk of the research on achievement and critical thinking skills has been done with college or graduate students, and little can be found with younger students.*

Critical Thinking and Gender

The relationship between gender and critical thinking is a topic of much debate and little conclusive evidence. The research is both mixed and contradictory (King, Wood, & Mines, 1990). Some studies indicate a relationship that favors males (Leach, 2011; Simon, 1974), others that favor females (Srinivasan & Cooks, 2005), and others that show no relationship at all (Ben-Chaim, et al., 2000; El Hassan & Madhum, 2007; P. Facione, 1990c; P. Facione, et al., 1995).

Though there is little empirical evidence to suggest differences in critical thinking skills by gender, dispositions may account for some of the visible differences between men and women. The leader in the gender bias argument, Thayer-Bacon (1993), argues that the conceptualization and instruction of critical thinking is biased toward males in that it is defined by "male attributes" and that "woman's logic" should be integrated into the definition. Consideration for personal experiences, emotion, and feelings should be

considered part of the critical thinking process since they guide actions. Wood (2012) also suggests that differences in communication style may be cause for perceived differences between the sexes. Women tend to build relationships through discussion and converse in a supportive and a feeling-centered manner. Men, on the other hand, are more competitive in their communication, and discussions are often a competition to prove a point of view (Wood, 2012). Thus, males may be perceived to have higher critical thinking skills

Additionally, Halpern (2004) suggests there are significant differences between the cognitive abilities of men and woman. Males, he says, tend to be better at spatial reasoning, and math and science standardized tests when they are not tied curriculum. Females, on the other hand, excel in writing and content area tasks that are tied to the curriculum. While there is little evidence to indicate a difference in critical thinking skill levels between men and women, there are differences in how each of the sexes approach critical thinking learning activities.

Critical Thinking and the Middle Years

Many psychologists argue that thinking skills are developmental, progressing with age. King and Kitchener (1994) developed the well known Reflective Judgment Model to illustrate the interdependence between cognitive ability and maturity. According to this model, reflective judgment does not adequately develop until adulthood. Various studies support this claim, indicating a strong correlation between critical thinking skills

and levels of cognitive development (Brabeck, 1983; P. Facione, 1990a; Kurfiss, 1988; Valiga, 1990). Critical thinking and reasoning, it can be argued, are developmental.

This is not to say, however, that middle school is not a crucial time for critical thinking and learning. Kay (2009), president of the Partnership for 21st Century Skills, argues that "combining proficiency in 21st century skills with core subject knowledge should be at the heart of middle school education" (p.43). Middle school students are ripe for critical thinking skills, among other 21st century skills, because they have a solid foundation on which to build. They are at an age when they are beginning to ask questions, think abstractly and explore the broader world (Kay, 2009). Middle school students tend to make great strides in their abilities to consider multiple perspectives and develop their own learning activities. Adolescent students are constantly examining themselves and questioning their worlds (*This we believe: Successful schools for young adolescents*, 2003). Furthermore, they are very social and enjoy opportunities to talk and collaborate with peers. The middle years may be the most important time to develop critical thinking skills by exploring values, assumptions, and basic principles.

The National Middle School Association (2003) recognizes that middle school students are greatly affected by the opinions of others and are prone to adopt the viewpoints of respected adults. Guiding students toward their own personal beliefs and values is a challenge the association sets for general middle school instruction. This aligns well with critical thinking instruction. Middle schools must foster this intellectual transition in students.

According to P. Facione, Gittens, N. Facione, & Winterhalter (2011), middle school is particularly important because it is a time when critical thinking skills can thrive or fail depending on the instruction. Middle school teachers serve a unique role, not only as teachers but also as role models and mentors. Middle school children are at an age where they seek to define themselves and look to adults for support and mentorship. Modeling critical thinking dispositions has been found to have the greatest impact on student dispositions, far more than direct instruction (Strahan, 1986). Eighth grade is a time when teachers might greatly impact student critical dispositions among other qualities.

While middles school students are not yet voting or serving as jurors, they do regularly use critical thinking. They need critical thinking to organize their time and assignments, to determine how not to lie, to avoid verbal and physical fights, to anticipate the demands of adults, to evaluate the suggestions of peers, and to explain issues to parents (P. Facione, et al., 2011). If they do not develop these skills in middle school, they are at a real disadvantage when they enter high school (Kay, 2009). Though middle school students are not at the peak of their critical thinking abilities, they do need instruction to develop their potential.

Critical Thinking and Culture

Confucian orientations toward learning. According to Hofstede, Hofstede, & Minkov (2010) humans are programmed with "software," or value patterns, which affect thinking, decisions, and how we work together. Pedagogy and traditions of learning are

deeply rooted in cultural history (Chan, Ho, & Ku, 2011; Turner, 2006). Though we all share universal human traits, a learned culture controls daily interactions and learning styles. Hofstede et al. (2010) assert that values define culture, and that these values are constant from one generation to the next. Though practices may change over time, the underlying values do not. "Software," as the call it, is an unconscious infrastructure, which creates the foundation for who we are. Cultural programming begins during infancy and is solidified throughout childhood. Hofstede et al. (2010) explain that students will continue to develop their mental programming in school with the influence of teachers and classmates who are also part of the shared culture. The combination of "software" along with environment will affect learning style and approaches to critical thinking (K. Brown, 1998).

Confucian heritage cultures show similarities in how they approach education. These Confucian cultures include Chinese, Taiwanese, Singaporean, Hong Kong, Japanese, and Koreans. These cultures tend to be hierarchical, have strong respect for teachers and their ability to impart knowledge, and believe success is the result of hard work. Redding (1980), a professor at Hong Kong University, discusses the Chinese learner and the influence of culture:

> The Chinese student, if he has been initially educated in his own culture, and his own language, will have begun to use a set of cognitive processes which give him a "fix" on the world in a very distinctive kind… It is possible to see some rationale for the noticeable tendency of Chinese to excel in certain subjects, particularly the applied sciences, where "the individual and the concrete" is paramount, and for

> their tendency not to move naturally into the abstract realms of philosophy and sociology… The most appealing explanations for it center upon differences in cognitive structures of a fundamental kind (as cited in Geert Hofstede, et al., 2010, p. 262).

In contrast, Western cultures tend to have a more fluid perspective on power, have a constructivist approach to knowledge, and believe aptitude is a large part of success (Geert Hofstede, et al., 2010). These cultural differences have influenced the use and conceptualization of critical thinking in the classroom. It is not to say that critical thinking is absent in educational settings that do not overtly encourage Western-style critical thinking opportunities such as classroom debate and the questioning of knowledge. Critical thinking may simply look different to the outside observer (Turner, 2006).

Reflective thinking is an integral part of learning in Confucian societies. In *The Analects* Confucius says that "learning without thought is labor lost; thought without learning is perilous" (as cited by Lee, 1996, p. 34). Rote learning, then, is a waste of time if thought and reflection do not accompany. Studying and learning require one to ponder, sift, and practice. Confucius himself taught using more of a Socratic method of questioning than one of lecture (Lee, 1996). This sense of learning through personal questioning and deep reflection, rather than reaction, is at the foundation of Chinese philosophy.

The contradiction between the stereotypical rote Chinese learner and Confucian reflective thinking can be explained by the words of Confucian scholar, Zhu Xi:

> generally speaking, in reading, we must first become intimately familiar with the text so that its words seem to come from our own mouths. We should then continue to reflect on it so that its ideas seem to come from our own minds. Only then can there be real understanding. Still, once our intimate reading of it and careful reflection on it have led to a clear understanding of it, we must continue to question. There might be additional progress. If we cease questioning, in the end there'll be no additional progress (as cited by Lee, 1996, p. 35).

Memorization, Zhu Xi argues, is an important part of the Confucian learning process in that it is the first step toward greater understanding. The role of knowledge in Confucian tradition, then, is viewed similarly to that of some Western philosophers and cognitive scientists. Once one has studied the words, one can reflect, ponder, and question, thus reaching higher levels of thinking. While rote learning is an important part of the Confucian learning process, it is a means to higher order, reflective thought.

Additionally, cultural differences in the conceptualization and interpretation of critical thinking may explain part of the contradiction. If critical thinking is defined as "purposeful self-reflective judgment," Chinese might perform well because reflective judgment is a focus of the Confucian tradition of learning (Biggs, 1996). If, instead, the definition emphasizes "challenging assumptions," then Chinese students might falter (Bond, 1996; Tiwari, et al., 2003). Chinese culture discourages some traits often associated with critical thinking, such as experimentation and public questioning (Turner, 2006). At the same time, Western students may be more adept at independent thinking, but not necessarily reflective judgment. Tiwari et all. (2003) found that Chinese are good

at identifying useful solutions, but less skillful at creating new solutions when one is unavailable[3]. A qualitative study by Turner (2006) found that Chinese students studying in UK had difficulty adjusting to a Western definition of critical thinking rather than the actual act of employing critical thinking skills. Turner indicates that that lower critical thinking performance may be due to dispositional factors and unclear thinking expectations rather than a skill deficiency. Expectations for critical thinking vary across cultures (Turner, 2006).

Misconceptions of Chinese students may also be due to a different expression of thought. It is uncommon for students in Confucian cultures to challenge a teacher or to debate a topic in class. Students are often quiet during class, listening to teachers reverently. At the end of class, however, it is not uncommon for Chinese students to speak privately with the teacher. At this time they may ask questions or engage in debate (Biggs, 1996). Additionally, students in collectivist societies often spend the bulk of their free time working together in study groups. Study groups are often a place for discussion and debate, thus promoting deeper thinking (Biggs, 1996). Reflective thinking, a conception of critical thinking, is deeply rooted in Confucian cultures, though it is likely revealed differently. Chinese students and Western students may also practice critical thinking skills differently. Western students tend to exercise their critical thinking abilities during classroom time, while Chinese students prefer less formal settings. Impressions of Chinese students as impassive learners may be an issue of style rather than substance.

[3] Note that experimentation was not part of the expert Delphi panel's definition described earlier in this paper. Both independent thinking and reflective judgment were.

Culture and critical thinking performance. A disconnect between perceptions of Asian learners and their performance on standardized tests is often referred to as the "Asian Paradox." We know Asian students tend to perform well on standardized tests, often outperforming American students. Korean, Japanese, Singaporean, and Hong Kong students regularly excel on the International Association for the Evaluation of Educational Achievement (IEA) tests in math and science (National Center for Education Statistics, 2013; Programme for International Student Assessment, 2012). The question, as it pertains to this study, is: How well do Confucian heritage students perform on critical thinking assessments and how well do they adapt to the demands of Western classrooms? The research in this area is limited. Recently, a couple of studies have looked at the relationship between critical thinking dispositions (the desire, rather than the ability, to think critically) and ethnicity. Two studies compared critical thinking dispositions of Hong Kong nursing students and found their scores averaged lower than test norms (Ip, et al., 2000; Tiwari, et al., 2003). Ku and Ho (2010) also found that Chinese students scored lower in critical thinking dispositions than their Western counterparts. Dispositions may be lower because of Confucian ideals of respect for elders, obedience, and conservative expression of emotions (Geert Hofstede, et al., 2010). The expression or attitude toward critical thinking may be lower than in Western societies, but little is known about skill level.

Another study by Watkins and Biggs (1996) compared thinking of students from Hong Kong, Malaysia, Nepal, Nigeria, Philippines, and Australia in terms of motivation and learning strategy. Motivation and learning strategies were divided into two levels - a

surface approach and a deep approach. Surface motivation and surface learning strategy are aimed at "getting by" and attaining the minimum standard. This is often by rote learning. Deep motivation and learning in this study were characterized by deeper intrinsic interest in a subject and the desire to make connections between ideas and diverse readings. Levels of motivation varied cross-culturally, with Australian, Nepalese, and Brunei students believing that a deeper approach to learning is necessary for success, while Hong Kong, Nigerian, and Indonesian had more surface levels of motivation. Approaches to learning, however, were dissimilar, presenting an interesting paradox. Australian students were less likely to use deeper learning strategies than Filipino, Hong Kong, and Malaysian students. Western students believed deeper learning was necessary, but their actual learning approaches were lower level. Learning approaches and learning styles may follow cultural lines.

Despite lower inclinations for critical thinking, research has found that with time Chinese students can adapt to new situations and expectations. Tiwari et al. (2003) and Turner (2006) both found that within a year Chinese students had adapted to the expectations for "deeper thinking and understanding" in Western classrooms. Volet and Renshaw (1996) found that Singaporean Chinese students initially struggled with learning style expectations, particularly that of deep understanding vs. memorization, but they were able to adapt to the expectations and were academically successful. Students learned to focus more on understanding main ideas rather than always seeking a correct answer or learning by memorization. Depending on the educational context, Chinese students have shown that they can adapt to new expectations for thinking and learning.

Thinking styles, then, may be more malleable and are likely affected by classroom expectations and school culture.

Summary

In summary, critical thinking is not a novel concept, but there is a growing need for focused discussion around its meaning, instruction, and assessment. Socrates asserted the importance of critical thinking in the pursuit of truth, Freire argued its necessity in obtaining personal freedom, Dewey viewed it as an educational ideal, and present day politicians and business leaders see it as an imperative for democratic and economic prosperity. How to define critical thinking has been a source of much debate, but recently the APA published a consensus definition for critical thinking as the "purposeful, self-regulatory judgment which results in interpretation, analysis, evaluation, and inference, as well as the explanation of the evidential, conceptual, methodological, criteriological, or contextual considerations upon which that judgment was based" (P. Facione, 1990d, p. 3). This common conceptualization may help drive instruction and assessment practices in the classrooms. While critical thinking instruction has been part of teacher training and educational discussions for some time, the assessment of it is less examined. We know it has a high correlation with grades and some standardized tests, but the details of those relationships remain vague. Critical thinking assessment is an essential part of the educational program. Assessment will drive further instruction, provide data for policy and program decisions, and create greater accountability. Critical thinking assessment is not, however, without its difficulties. Some have argued it to have

gender bias, age bias, and ethnic bias. Such issues will be examined in some form during this study. Foremost, however, this study will investigate the extent to which critical thinking assessments align with current reporting systems in some schools – the MAP test and grades. Additionally, by understanding demographic differences by gender, school curriculum, and culture we can further understand the relationship critical thinking plays in different settings and with different demographics.

CHAPTER 3: METHODOLOGY

"He who asks a question is a fool for five minutes; he who does not ask a question remains a fool forever."
– Chinese Proverb

Statement of Study Purpose

The purpose of this study is to examine the relationship between critical thinking skills and academic performance of middle school students at an American international school in China, and to determine the degree to which student demographic characteristics moderate the relationship.

Research Questions

1) What is the level of critical thinking, as measured by the California Critical Thinking Skills Test, of eighth grade students at an American international school in Shanghai, China?

2) To what extent is there a relationship between critical thinking and student achievement, as measured by grades and Measures of Academic Progress (MAP) test scores?

3) Do student levels of critical thinking vary in terms of the following demographic variables?

- Gender
- Tenure at Shanghai American School
- Native language

Methodology and Rational

This quantitative study examines the relationship between critical thinking skills, as measured by the California Critical Thinking Skills Test (CCTST), and performance on the MAP test and course performance as indicated by grades. The analysis also investigates whether there is a relationship between critical thinking skills and gender, tenure at SAS, and native language. The quantitative analysis identifies the existence and extent of any relationships.

A one-way ANOVA to compare means of demographic variables with critical thinking, grades, and MAP Rausch Unit Scores (RIT) scores is the first stage of analysis. Initial results help determine how variables are entered for the subsequent stepwise analysis. A stepwise regression is then used to determine the strongest predictor of critical thinking. The researcher hypothesized that MAP test scores would be the most important variable to influence critical thinking. It has been established in the literature that critical thinking skills are linked to high academic achievement (Cabrera, 1992; Garett & Wulf, 1978; Steward & Al-Abdulla, 1989; Williams, 2003; Williams & Stockdale, 2003). There is also evidence, however, that low critical thinkers can find adaptive strategies in order to perform moderately in academic contexts (Tiwari et al., 2003; Turner, 2006; Volet & Renshaw, 1996). Thus, the researcher also hypothesized that low critical thinking scores would have a moderate correlation with grades – reflecting adaptive strategies for classroom success. Due to the contradictory evidence between gender, ethnicity, and critical thinking, no relationship was predicted between gender, native language, and critical thinking skills.

In this context, native language serves as a proxy for culture. Identified languages are grouped into two broad ethnic groups: Anglo for English speakers and Confucian for Cantonese, Mandarin, Korean, Japanese, and Taiwanese speakers (Geert Hofstede, et al., 2010). Languages not falling in either of these two categories were excluded from this part of the analysis.

Language can be viewed as an artifact of a culture. Hofstede (1983) defines culture as "collective mental programming: it is the part of our conditioning that we share with other members of our nation, region, or group but not with members of other nations, regions, or groups" (p. 76). Experiences as a member of a group or society help one to identify with a particular culture.

The research, as discussed in the literature review, shows the cultural influence on critical thinking dispositions (Ip, et al., 2000; Ku & Ho, 2010; Tiwari, et al., 2003), but there is no conclusive evidence in terms of skill levels. P. Facione does suggest, however, that critical thinking skills and dispositions interact with one another and are mutually reinforcing (P. Facione, 2000), but this has not yet been tested with cultural variables. The expression of critical thinking may be culturally bound (Turner, 2006), but its influence on critical thinking skills test performance lacks substantial testing as of date (Neisser, et al., 1996; Zhang, 2003).

The research regarding the relationship between gender and critical thinking is also mixed. Some studies indicate a relationship that favors males (Simon, 1974), others that favor females (Srinivasan & Cooks, 2005), and others that show no relationship at all (Ben-Chaim, et al., 2000; El Hassan & Madhum, 2007; P. Facione, 1990c; P. Facione, et

al., 1995). The hypothesis, therefore, was that gender does not predict critical thinking skill levels.

Tenure at Shanghai American School was also predicted to correlate with thinking skills. The literature reveals a disadvantage for Chinese students who move abroad and study in Western institutions, though with time they were able to adapt to new expectations and modes of thinking (Turner, 2006). Therefore, students new to Shanghai American School in the 2012-2013 academic year may need more time to adjust to new expectations and were predicted to have lower overall scores on the CCTST.

As defined by this study, critical thinking is "purposeful, self-regulatory judgment which result in interpretation, analysis, evaluation, and inference, as well as the explanation of the evidential, conceptual, methodological, criteriological, or contextual considerations upon which judgment is was based" (P. Facione, 1990d, p. 3). This definition is based on the conceptual understanding of critical thinking as defined by the expert panel in the Delphi Report and published by the APA. Though dispositions are an important element of critical thinking, this study focuses solely on critical thinking skills.

Population and Sample

314 8th grade students from both the Puxi and Pudong campuses of Shanghai American School make up the population of this study. Students are generally wealthy and high performing. Approximately 98% of students are predicted to take at least one AP or IB course in high school, and 99% of students will likely go to college or university upon graduation (McElroy, October 3, 2012). According to the Challenge

Index, a measurement first proposed by Jay Matthews of the Washington Post to measure challenge based on the number of AP and IB exams taken at a school, Shanghai American School would rank number one if compared to private schools in the United States. Middle school students will likely go on to take challenging high school courses.

As discussed in the literature review, middle school students are ripe for developing critical thinking skills, and the middle years are essential for developing thinking skills before they encounter more rigorous courses in high school. Middle school students are at an age when they are beginning to ask questions, think abstractly and explore the broader world (Kay, 2009). Additionally, 8th grade students have sufficient language, knowledge, and reading backgrounds for this test.

Shanghai American School was selected because of its reputation for offering a high quality American education and for the inclusion of critical thinking as part of the mission and values. Shanghai American School also serves a diverse population of students from various cultures.

This study samples nearly the entire population of eighth graders at Shanghai American School. The sample consists of 297 of the 314 8th graders at SAS. 3 students were dropped due to computer failure, 1 student requested to opt out of the assessment, and 13 students were absent on test taking day. Though students who spend less than 15 minutes on the test are typically discarded as false tests, they are included in the final analysis because computer logout issues required a number or students to login a second time and skip through half the test, invalidating their test time.

There is a relatively even spread of demographic features within this sample. The sample includes 150 (50.5%) males and 147 (49.5%) females. 142 (47.8%) students in this study identify a native language categorized under the Anglo cultural cluster, and 140 (47.1%) identify a language falling under the Confucian cluster. 15 (5.1%) students identified a native language that did not fall under either of these two cultural clusters and were dropped from this part of the analysis. Students were asked to represent their tenure at SAS by selecting one of the following categories: 0-1 years, 2-4 years, 5-8 years, 9+ years. 26 (8.8%) students indicated that they had attended SAS for 0-1 years. 80 (26.9%) students attended between 2-4 years, 97 (32.7%) students 5-8 years, and 94 (31.6%) students had attended for 9 or more years. The mean number of years was 5.84. For analysis purposes, the mean of each tenure category was used.

Critical Thinking Skills Test

Instrument. This study requires the collection of three data sets. The first is the critical thinking skill levels of 8th grade students. This is measured by The California Critical Thinking Skills Test M-25 (See Appendix A for sample questions). The CCTST family of tests are all based on the Delphi consensus definition and the definition used by this study (P. Facione, 1990d). The CCTST-M series is designed for students in grades 6-9 or for individuals of similar reading levels. The test is based on the original CCTST test, which is used for undergraduate students and above. Test questions are adjusted to appropriate reading levels, to engage younger audiences, and for developmental

appropriateness. Instruments in the CCTST family have been used by school districts, program coordinators, and curriculum specialists globally (August, July 31, 2012).

Test design. The CCTST-M25 measures core cognitive skills associated with critical thinking using everyday contexts. Questions are not subject matter specific and require no prior content knowledge. Seven subsets or scales associated with critical thinking are covered by the test, though test makers urge users against using scaled scores as independent indicators. These scales, as identified by the Delphi Report, are: 1) Analysis and Interpretation; 2) Inference; 3) Evaluation and Explanation; 4) Deductive Reasoning; 6) Inductive Reasoning; 7) Numeracy. Numeracy was added most recently to assess recognition and understanding of qualitative information as represented by graphs, charts, tables, and diagrams.

The CCTST-M25 asks students to analyze and interpret 25 charts, tables, and primary texts. Students must select the best answer among five answer choices. There is one correct answer among four distracter items. The questions do not test factual knowledge and are about everyday situations of which students should be familiar.

This study also includes an additional section for personal demographic information (See Appendix B). Students were asked to identify their gender, years at SAS, transfer school if applicable, native language, their passport country, mother's passport country, and father's passport country. Information gathered in this section was used to analyze factors contributing to critical thinking skills.

Upon completion of the test, scores for each of the subsets as well as overall critical thinking scores were made immediately available to teachers. Teachers were

given the option of sharing results with students. Scores are reported in two ways: as a score along a continuum, which ranges from not manifested to strong, and as percentile score. Percentile norm comparisons were also made available.

Administration. The test was administered to the sample of 297 students over four days. During the social studies block, each class of students was given 45-minutes to complete the CCTST-M25 test in the SAS computer lab. Test data was saved directly to the testing system and secured with a unique administrator ID and password.

Scoring. The CCTST generates a total score as well as scores for each of the subsets. While each scale score can inform curricular and instructional decisions, they should not be used as independent factors. Each of the scales interact with one another to aid reflective judgment and critical thinking (P. Facione, 1990a).

A *superior* score between 84-100 indicates an excellent critical thinker who can independently solve complex reasoning problems and complete high level learning tasks. *Strong* scores of 75-84 identify individuals with strong thinking skills who would benefit from an integrated approach to critical thinking. With instruction, they can likely improve their critical thinking skills in various situations or contexts. A score of 66-74 identifies students with *emerging* critical thinking skills. These individuals will require basic interventions to fully benefit from standard curriculum and instruction. Finally, a *not manifested* score of 65 and below signifies a language barrier, lack of effort, or substantial critical thinking limitations.

Psychometrics. Content validity for the CCTST instrument and the reported subsets of critical thinking skills are supported by the results of the Delphi consensus

expert study of critical thinking. Construct validity is supported by numerous peer reviewed publications on correlations with other instruments purported to measure critical thinking skills as well as student CCTST gains after completing courses focused on critical thinking (Barak, Ben-Chaim, & Zoller, 2006; "CCTST M- Series test manual," 2013; Sullivan-Mann, Perron, & Fellner, 2009; Yang, Ya-Ting, Newby, & Bill, 2008). Criterion validity, the ability of the test to predict a criterion behavior external to the test itself, is also supported by numerous peer reviewed publications (Denial, 2008; Giddens & Gloeckner, 2005; Vendrely, 2007; K. Williams, et al., 2003).

Validation studies for the M-Series instruments were conducted in public and private educational settings across the United States and percentiles for comparison are derived from these studies (August, July 31, 2012). The internal consistency coefficient for a dichotomously scored instrument is the Kuder Richardson-20. KR-20's for the M-Series validation samples ranged from .78-.82. Any KR-20 above .70 is considered strong given that the instrument also provides scale scores for a number of different constructs ("CCTST M- Series test manual," 2013).

Reading levels and the relevance of the item topics were confirmed with grade school teachers ("CCTST M- Series test manual," 2013). Individual norms for each grade level have not been published at this time, and the publishers maintain that there is insufficient evidence to assume each grade level is a significant predictor of mean scores within its intended test taker group. The norms calculated for the CCTST-M25 test are based on the scores of students in grades 6-9 ("CCTST M- Series test manual," 2013).

Measures of Academic Progress (MAP)

The MAP is a computerized adaptive test (CAP) published by the Northwest Evaluation Association (NWEA). At Shanghai American School, three MAP tests are used: Reading, Language Usage, and Math. Each test is customized to adjust to the skill level of the test taker, testing until equal precision is reached. Test questions pull from a bank of over 70,000 questions aligned to specific state content standards. Computer adaptive tests, such as the MAP, have better validity than traditional paper-pencil tests (Way, 2006) because they are able to zero in on an accurate result at all proficiency levels (Van Horn, 2003). These tests will question students until they answer 50% of the questions correct and 50% incorrect. The MAP test is also unique in that it reports percentile scores, achievement scores, and Rausch Unit Scores (RIT). The RIT score is an equal-interval scale, like feet and inches, which is independent of grade level. The mean of the three RIT scores is used in the analysis.

Over 800,00 students have taken the MAP in the Unites States, as well as in international settings (Van Horn, 2003). Construct validity tests found MAP tests to be "well defined, proved to be unidimensional equivalent across grades, and have the same patterns across academic years" (Wang, McCall, Jiao, & Harris, 2012, p. 7). Tests have marginal reliabilities in the low to mid .90's (Northwest Evaluation Association, 2003). The MAP has been used not only to provide school stakeholders with formative assessment data on student learning, but also to offer correlation data to inform educational research. The MAP was used, for example, to measure consistency between state expectations for proficiency (Dahlin, Xiang, Durant, & Cronin, 2010), assess

learning growth rate factors between schools (Xiang & Hauser, 2010), and in connection with the ACT test to measure college readiness (Northwest Evaluation Association, 2011). The MAP test is a widely used and validated test of student learning.

Each year, SAS administers the Reading, Language Usage, and Math MAP tests in the fall and spring for grades 3-8. The 8th grade Spring data was for this study because it was administered at relatively the same time as the CCTST. All student MAP data is logged by an identification number in the NWEA website and can be downloaded using a secure username and password provided by SAS.

Grades

Student academic achievement is reported as a percentage score based on demonstrated mastery of academic standards. The SAS assessment manual states that "final grades will be derived from accurate assessment information, gathered through a variety of assessment methods that are appropriate and relevant to learning outcomes, standards and benchmarks" (Shanghai American School, 2011b, p. 11). Only core course percentages from social studies, language arts, math, and science were included for analysis. Dispositions toward learning, such as cooperation, preparation, integrity, and effort should not be included in academic scores, as they are represented on a separate learner profile, which is not used in this study. Final academic percentages are assigned using teacher discretion but should represent achievement of curricular standards and benchmarks as well as enduring knowledge and skills. Standards are based on the Common Core State Standards.

Student academic data was retrieved by student ID from Power School, a school-wide database for reporting and storing student information. The mean of percentages from core academic classes was used to represent grades for the analysis. Percentages represent the course grades for the 2012-2013 academic year and do not include previous years.

Demographics

Demographic data was obtained using two methods. Students self-reported demographic information as part of the critical thinking skills test. Their answers were cross-referenced with demographic information provided by parents and stored on the school's secure database, PowerSchool. In cases of discrepancy, further inquiry with students, counselors, and parents was required.

Procedures

The research study, its rationale, and procedures for data collection were introduced to social studies faculty via personal meeting. Teachers were asked to introduce the study to students and parents, distribute parent consent forms via email, and attend the scheduled test time with their students. On test day, the administrator used a pre-planned script to explain the study to students, its rationale, and procedures. Students were asked to sign an assent form if they agreed to participate. In the event they wanted to opt out, the option of reading in the library during the test time was offered. Students were told that their teachers would have access to the critical thinking test results, and

that they could request access to their scores upon completion. It was made clear, however, that the test would not impact student grades. Though the testing system has the ability to report scores directly to students at the end of the test, this method of sharing results with teachers was chosen because of teacher concerns about student competition. At the same time, students were motivated by knowing they could view their scores at a later time, and that the teacher had access to scores.

Analysis

The first step of analysis examines bivariate correlations among all pairs of variables - critical thinking skills, MAP/RIT scores, grades, gender, tenure at SAS, and native language. In a second stage, multiple regression analysis uses Pearson's Correlation Coefficient to assess the relationships between grades, MAP test RIT scores, and critical thinking skills. A part of the multiple regression analysis also considers whether demographic variables, gender, culture, and longevity, impact critical thinking.

CHAPTER 4: RESULTS

"Insanity is doing the same thing over and over again while expecting a different outcome"
- Albert Einstein

Introduction

The purpose of this study is to measure critical thinking skill levels of 8th grade students at Shanghai American School, and to determine the relationship between critical thinking skill levels and other measures of academic achievement. Core course grades and Measures of Academic Progress (MAP) test scores are used to represent academic achievement in this study. The secondary purpose is to identify demographic variables, gender, culture, and tenure at the school that may moderate this relationship.

Descriptive results are presented first, providing an overview of the critical thinking skill levels of students and relative strengths and weaknesses in each of the subsets for critical thinking – analysis, inference, evaluation, induction, deduction, and numeracy. Descriptive data on student grades and MAP test scores are included to provide insight into the academic achievement level of the sample. The demographic section also includes one-way ANOVA tests on how achievement variables – MAP Rausch Unit Scores (RIT), grades, and critical thinking - vary by demographic variables. The second section of this chapter introduces results of the stepwise regression analysis for all independent variables - grades, MAP RIT, gender, native language, and tenure.

Descriptive Results

Critical thinking overall. Descriptive statistics address research question #1: *What is the level of critical thinking, as measured by the California Critical Thinking Skills Test, of eighth grade students at an American international school in Shanghai, China?* Descriptive results are displayed in Table 1 and the histogram in Figure 1.

The overall score best represents general critical thinking ability. The overall mean score for respondents is 83.61, above the U.S. mean of 78.7. As compared to U.S normative data from grades 6-9, the average SAS 8th grader scored in the 69th percentile. The data shows a standard deviation of 6.5, a range from 60-97, with the histogram skewed to the left.

Table 1: *CCTST Overall Descriptive Statistics*

N	*Sample Mean*	*U.S. Mean*	*St. Dev.*	*Minimum*	*Q1*	*Median*	*Q3*	*Maximum*
298	83.61	78.7	6.5	60	80	84	83	97

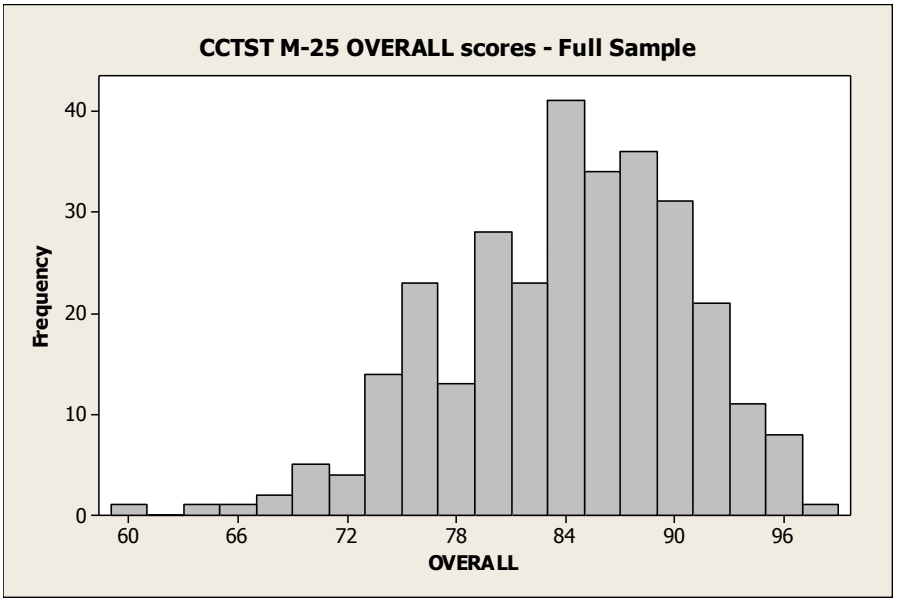

Figure 1: CCTST Overall Scores

Critical thinking test scores can be grouped into four performance assessment levels – superior, strong, emerging, and not manifested. Table 2 displays performance assessment levels for sampled students. A *Superior* descriptor indicates that a student scored 85 or above and outperformed the vast majority of test takers in the United States. 142 students in this sample fall under the *Superior* range. 92 students and the overall mean fall between 79-84, in the *Strong* range. 50 students score in the *Emerging* range, between 73-78. Critical thinking skills are *Not Manifested* when they are between 60-72. 14 students score as *Not Manifested,* indicating that they have either very low critical thinking skills or a false test from lack of effort.

Table 2: *Recommended Performance Assessments for the CCTST M-Series Overall Score*

	Not Manifested	***Emerging***	***Strong***	***Superior***
CCTST M-Series Overall Score	60-72	73-78	79-84	85 or higher
Number of SAS Students Scoring in Range	14	50	92	142

Critical thinking scaled scores. Though critical thinking is a holistic process, U.S. normative data shows that individuals and groups have relative strengths and weaknesses in each of the subsets: Analysis, Inference, Evaluation, Induction, Deduction, and Numeracy. Induction skills had the highest mean (87.91), while Deduction had the lowest (80.55). The following is the ranking order from strongest scale score to least manifested: *Induction* (87.91), *Inference* (85.34), *Evaluation* (82.99), *Analysis* (82.56), *Numeracy* (81.82), *Deduction* (80.55). Table 3 and Figure 2 represent results of scaled scores.

Table 3: *CCTST Descriptive Statistics: Analysis, Inference, Evaluation, Induction, Deduction, Numeracy*

Variable	*N*	*Mean*	*St. Dev.*	*Minimum*	*Q1*	*Median*	*Q3*	*Maximum*
Analysis	298	82.56	7.0	60	78	84	88	98
Inference	298	85.34	8.2	60	80	86	91	100
Evaluation	298	82.99	6.7	60	79	84	88	95
Induction	298	87.80	7.8	60	84	88	92	100
Deduction	298	80.55	6.9	60	76	81	84	97
Numeracy	298	81.82	7.7	60	78	82	89	96

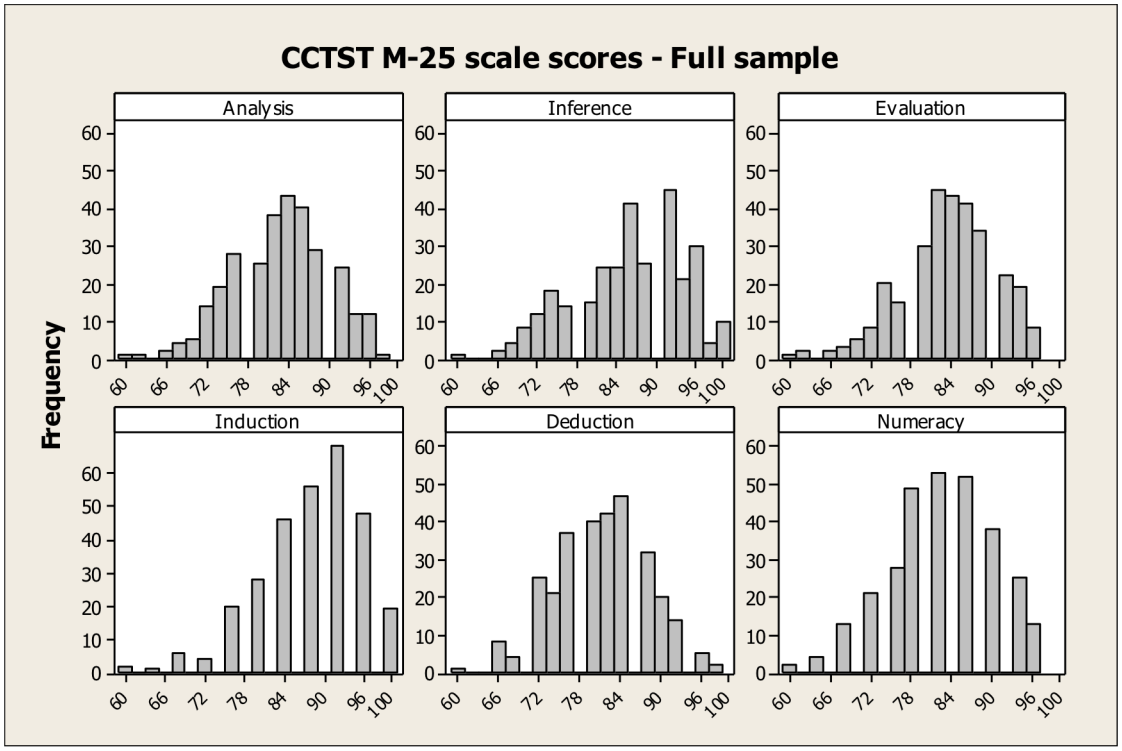

Figure 2: CCTST Scaled Score: Analysis, Inference, Evaluation, Induction, Deduction, Numeracy

Grades. Grades are represented as a percentage out of 100. The mean grade for each student was calculated with core academic classes: English, social studies, science, and math. Table 4 shows the mean grade percentage of all students is 89.33%, with a

letter grade equivalent of a high B+. Grade averages range from 64.13% (D-) to 98% (A) with a standard deviation of 5.63%. Figure 3 shows the histogram for mean grade score in the combined core subject areas.

Table 4: *Mean Grade Score Descriptive Statistic*

	N	***Minimum***	***Maximum***	***Mean***	***Std. Deviation***
GradeScore	297	64.1250	98.000	89.329	5.628

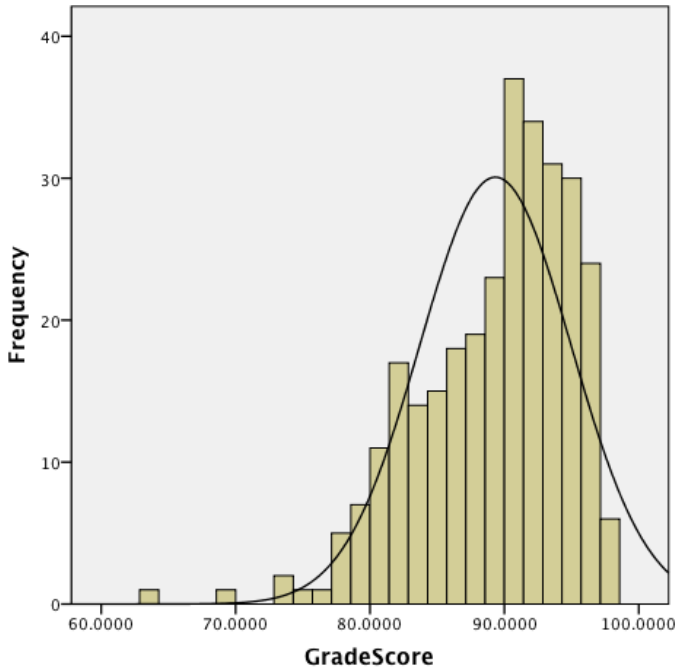

Figure 3: Mean Grade Score

MAP RIT scores. Table 5 and Figure 3, represent the descriptive data for MAP RIT scores. The mean RIT score, 242.45, is categorized as "high" as compared to all MAP test takers. Overall scores range from 217.67-263.67. When broken down by each of the three subsets (Reading RIT, Language Usage RIT, Math RIT), students were strongest in Math (92nd percentile). Language Usage (90th percentile) was slightly lower, and though still strong, Reading (84th percentile) was ranked lowest of the three.

Table 5: *MAP RIT Score Descriptive Statistics*

Variable	*N*	*Mean*	*U.S. Percentile Rank*	*St. Dev.*	*Minimum*	*Maximum*
Mean MAP/RIT	288	242.45		8.99	217.67	263.67
Reading MAP/RIT	288	235.46	84%	9.72	206	258
Language Usage MAP/RIT	290	234.73	90%	8.47	212	258
Math MAP/RIT	289	257.28	92%	12.92	218	300

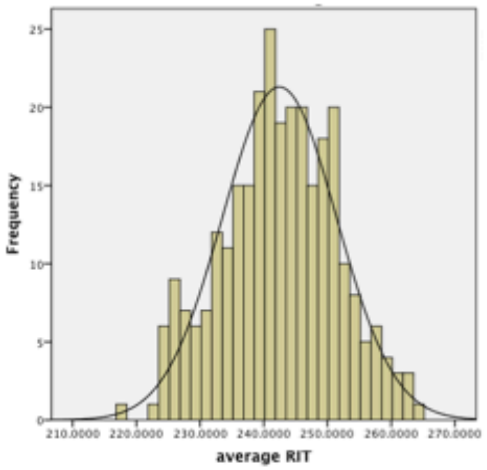

Figure 4: Overall Average RIT Score

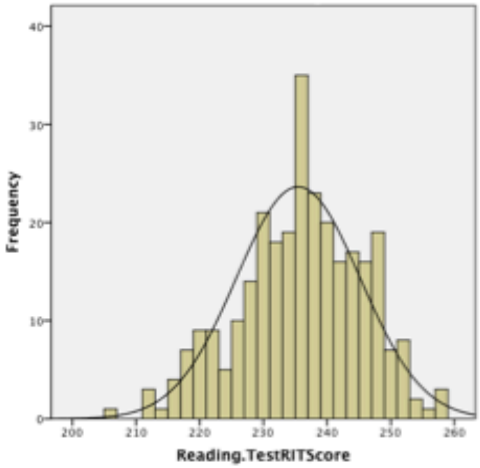

Figure 5: Reading RIT Scores

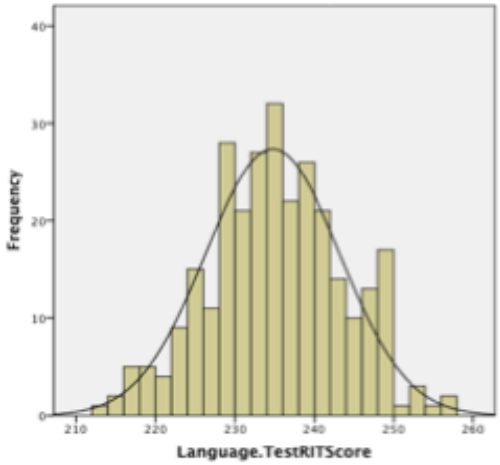

Figure 5: Language RIT Score

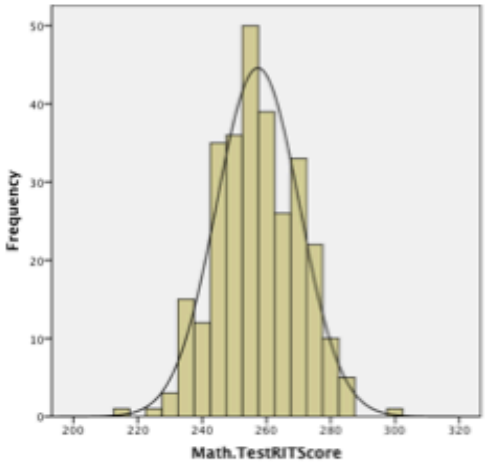

Figure 7: Math RIT Score

Figure 4: MAP/RIT Scores

One-way ANOVA tests. One-way ANOVA test were conducted to determine if critical thinking, MAP/ RIT scores, and grades vary by demographic variables. Demographic variables – gender, language, and tenure – were all considered, though tenure did not yield significant results (See Appendix E). Table 6 indicates that females have higher grades and higher Language MAP/RIT scores than males. Table 7 indicates that Confucian language is correlated with higher Math MAP/RIT scores. Results did not, however, indicate a significant relationship between critical thinking and any demographic variables.

Table 6: *Gender (female) on Critical Thinking, Grades, and MAP/RIT Scores*

Dependent		***Mean Square***	***F***	***Sig***
Critical Thinking	Between Groups	73.738	1.772	.184
	Within Groups	41.617		
Grades	Between Groups	860.483	29.763	.000
	Within Groups	28.912		
Mean RIT	Between Groups	95.963	1.188	.277
	Within Groups	80.745		
Reading RIT	Between Groups	190.010	2.018	.157
	Within Groups	94.152		
Math RIT	Between Groups	196.283	1.176	.279
	Within Groups	166.857		
Language Usage RIT	Between Groups	905.645	13.199	.000
	Within Groups	68.614		

Table 7: *Confucian Language on Critical Thinking, Grades, and RIT Scores*

Dependent		***Mean Square***	***F***	***Sig.***
Critical Thinking	Between Groups	2.127	.051	.822
	Within Groups	41.888		
Grades	Between Groups	26.109	.821	.366
	Within Groups	31.790		
Mean RIT	Between Groups	114.096	1.383	.241
	Within Groups	82.521		
Reading RIT	Between Groups	279.405	2.914	.089
	Within Groups	95.873		
Math RIT	Between Groups	3077.405	19.617	.000
	Within Groups	156.876		
Language Usage RIT	Between Groups	23.318	.316	.574
	Within Groups	73.790		

Regression Analysis

Three stepwise regression equations were used in order to address research questions 2 and 3: *To what extent is there a relationship between critical thinking and student achievement, as measured by grades and MAP test scores? Do student levels of critical thinking vary in terms of the following demographic variables: gender, tenure at Shanghai American School, and native language?*

Table 8 represents the first regression equation, using critical thinking as the dependent variable with two steps of independent variables. The first step includes all demographic variables, gender, language, and tenure, and the second step adds grades. The first step suggests no relationship between critical thinking and gender, language, or tenure at SAS. The addition of grades in the second step has a very significant effect. Grades are a strong predictor, and the percentage of variance explained (R2) increases dramatically from .006 to .349. Including grades does not change the relative importance of language and tenure, but the relationship between gender and critical thinking shifts from insignificant and positive to marginally significant and negative, indicating that females get slightly higher grades than males.

Table 8: *Regression of Critical Thinking on Demographic Variables and Grades (N=282)*

	Predictors	***Standardized Coefficients-Beta***	***t***	***Sig.***	***R^2***
1	**(Constant)**		82.841	.000	
	Gender	.075	1.256	.210	
	Language	.032	.484	.629	
	Tenure at SAS	.000	.004	.997	
					.006
	F = .586* **				
2	**(Constant)**		4.004	.000	
	Gender	-.103	-2.015	.045	
	Language	-.016	-.311	.756	
	Tenure at SAS	.007	.139	.889	
	Grades	.612	12.077	.000	
					.349
	F change = 145.864***				

The second regression, found in Table 9, includes mean MAP/RIT as an independent variable and third step. As discussed previously, gender is slightly significant with grades and language MAP/RIT, but then it moves back to being insignificant when mean MAP/RIT is added. Females may get slightly higher grades than males, but their academic achievement levels, as measured by the composite MAP/RIT test score, are relatively equal. The second step suggests that among those students with the same grades, females do less well on the critical thinking test, but this disappears when you take their math performance into account. Language and tenure are insignificant regardless of the additional independent variables. Grades and mean MAP/RIT are clearly significant, with mean RIT being the most highly correlated variable (β=-.612 with .000 significance). When mean MAP/RIT is included, the R2 moves from .351 to .531. Mean MAP/RIT is an even better predictor of critical thinking than grades.

Table 9: *Regression of Critical Thinking on Demographic Variables, Grades, and Mean MAP/RIT Score (N=274)*

Predictors		***Standardized Coefficients-Beta***	***t***	***Sig.***	***R^2***
1	(Constant)		80.166	.000	
	Gender	.066	1.084	.279	
	Language	.022	.325	.745	
	Tenure at SAS	-.004	-.053	.958	
					.005
	F = .411***				
2	(Constant)		3.873	.000	
	Gender	-.113	-2.199	.029	
	Language	-.024	-.445	.657	
	Tenure at SAS	000	-.005	.996	
	Grades	.616	11.976	.000	
					.351
	F change = 143.422 ***				
3	(Constant)		-5.328	.000	
	Gender	-.018	-.392	.695	
	Language	-.038	-.826	.409	
	Tenure at SAS	.009	.196	.845	
	Grades	.146	2.300	.022	
	Mean RIT	.620	10.143	.000	
					.531
	F change = 102.876***				

The final model, found in Table 10, takes a deeper look at MAP/RIT scores and uses the math subset rather than the mean MAP/RIT, thus excluding language usage and

reading.[4] Including math MAP/RIT, as opposed to the mean MAP/RIT, reveals a negative effect (β=-.114 with .024 significance), indicating that the Confucian language impact on critical thinking is largely an artifact of math performance. When math performance is factored out, non-Asians have slightly higher critical thinking scores. Confucian students perform better in math, but among those students who perform similarly in math, non-Confucian score slightly higher in critical thinking.

Table 10: *Regression of Critical Thinking on Demographic Variables, Grades, and Math MAP/RIT Score (N=275)*

Predictors		***Standardized Coefficients-Beta***	***t***	***Sig.***	***R^2***
1	(Constant)		80.633	.000	
	Gender	..067	1.098	.273	
	Language	.020	.307	.759	
	Tenure at SAS	-.002	-.032	.974	
					.005
	F = .419***				
2	(Constant)		-1.244	.215	
	Gender	.005	.105	.917	
	Language	-.114	-2.278	.024	
	Tenure at SAS	-.009	-.188	.851	
	Grades	.293	4.704	.000	
	Math RIT	.477	7.737	.000	
					.468
	F change = 117.265 ***				

[4] Language usage and reading were also analyzed but did not impact critical thinking. Results confirm that the reading and language effect is based on math performance, not only other areas of cognitive achievement (See Appendix D & E).

CHAPTER 5: DISCUSSION AND CONCLUSION

"A mind stretched by a new idea never goes back to its original dimensions."
- Oliver Wendell Holmes

Student learning data is an important tool for all educational stakeholders. A few of the many benefits of using high quality assessments of student learning are that they can provide feedback on school initiatives, serve as leverage for additional initiatives to increase student achievement, and guide programmatic decisions. Many argue, however, that a focus on assessment and a data driven educational culture has narrowed instruction and emphasized low-level thinking over higher order cognitive tasks which would require critical thinking, problem solving, and creative discovery (Conley, 2007; L. Resnick & Zurawsky, 2005). Essential in addressing the assessment debate is a discussion of the type of assessment and the extent to which tests measure a range of student skills, and the careful monitoring of unintended consequences (both positive and negative). Ultimately, how students are assessed will set the roadmap for teacher instruction and student learning.

Assessment of critical thinking skills is an educational priority. This study seeks to address this demand by investigating the relationship between critical thinking skills, Measures of Academic Progress (MAP) test scores, and grades, which can help educators identify the extent to which critical thinking assessment is embedded in established assessment practice. Secondly, it looks at demographic factors – gender, culture, and years of schooling at Shanghai American School – to provide further insight into critical thinking skills instruction and how educators can differentiate for a diverse student body.

This study seeks to influence discussions about higher order thinking as part of an explicit agenda for schools by focusing on critical thinking assessment. Student critical thinking skill data provides valuable information to teachers and school leaders as they guide instruction, design programs, and develop curriculum. Beyond implications that are specific to Shanghai American School, results add to the current body of research on critical thinking and suggest areas for further investigation.

Results of this study found that students at Shanghai American School are very strong critical thinkers, and that critical thinking scores are highly correlated with other measures of academic success – both grades and MAP scores. High correlations suggest that the grading policy and practices at Shanghai American School include critical thinking skills, thus setting a high standard for thinking and learning. MAP test score correlations also indicate that the MAP test, though intended only as an academic measure, also predicts critical thinking skill levels of students and can inform a thinking curriculum. A relationship between critical thinking and demographic variables such as gender, culture, and student tenure at the school is not reflected in the data. While the results of this study may not generalizable to other settings, the study suggests that studying critical thinking, as contrasted with standardized test results, may illuminate issues related to the development of a more rigorous curriculum that prepares students for the demands of the 21st century.

Implications and Future Research

School-level implications. Critical thinking skills assessment is an important first step toward thinking instruction and curriculum. Results generate discussions about

values, curriculum, and pedagogy to foster critical thinking. The data can also be used to support individual students in course selection and to determine their own learning objectives. Finally, the data helps generate discussions with stakeholders (students, parents, board members) about educational values and desired outcomes.

This study takes this first step by determining the critical skill level of students and thus leading the way to a broader discussion for critical thinking teaching and learning. In the case of Shanghai American School, students are overall strong critical thinkers, scoring higher than 69% of all test takers. Scores are grouped into four performance assessment levels – *superior, strong, emerging*, and *not manifested*. Though the mean score falls in the *strong* category, 142 of the 297 test takers score as *superior*, the highest assessment level. This strong overall score indicates that students are able to maintain focus and apply core critical thinking skills – analysis, interpretation, inference, evaluation, explanation, induction, deduction, and numeracy – to a variety of real-world scenarios. Strong critical thinking skills are likely indicative of a competitive admissions process, effective instructional practices, and a thinking curriculum at SAS. An analysis of factors contributing to strong thinking skills is beyond the scope of this study, but would be an important next step for continuing and replicating strong results.

The SAS mission and core values support a well-rounded education that will empower students with the "creativity, critical thinking and lifelong passions for learning [that] are essential to personal fulfillment and to meet the challenges of the future" (Shanghai American School, 2012, para. 2). Overall results suggest SAS is in at least partial fulfillment of its mission and that learning outcomes are aligned to core values.

These findings provide valuable information to the Board, parents, teachers, and students. The superintendent and board members could, for example, use results as another data point to demonstrate student learning to stakeholders and accreditation bodies. Human Resources might include results in marketing brochures to illustrate the value of an SAS education. Principals and teachers could incorporate data as a way to track the development of student learning over time and hone in on instructional practices that specifically support critical thinking development.

Student-level critical thinking data, along with other information sources, can also help counselors guide students to appropriate high school courses. Critical thinking skills are a predictor of academic success and can help determine if students are prepared for rigorous high school courses like the International Baccalaureate and Advanced Placement (P. Facione, 1990a; R. Williams, et al., 2003). Students whose scores indicate emerging thinking skills, may instead be recommended for courses that can support critical thinking development. Because this particular study is for research purposes and subjects are minors, names remain anonymous and student level analysis cannot be done with the current data set. Schools would benefit, however, from administering their own thinking skills assessment and using the data to guide individual counseling and instruction.

The early teens are a point of intellectual transition for students and critical thinking skills data can inform appropriate instruction (P. Facione, et al., 2011; Kay, 2009; *This we believe: Successful schools for young adolescents*, 2003). Critical thinking

matters and by assessing, analyzing, and planning action steps based on thinking data, schools like Shanghai American School champion their values.

Grading policy implications and future research. Little to no data is available on the relationship between critical thinking skills and academic achievement of younger students, particularly those in the middle school years. However, numerous college and graduate level studies have highlighted a strong relationship between critical thinking and academic achievement, as measured by grade point average (GPA) (Cabrera, 1992; Garett & Wulf, 1978; Steward & Al-Abdulla, 1989; K. Williams, et al., 2003; R. Williams, et al., 2003). This study supports the link between academic achievement and grades when schools have invested in sound grading practices.

The results of this study show a strong correlation between critical thinking skills and grades. Grades and critical thinking skills are both highly significant dependent and independent variables, meaning students with high grades generally have high critical thinking scores and visa versa. Additionally, grades are also highly correlated with MAP scores. If grades are supposed to reflect the skill level of students, they are accurately doing so.

Critical thinking tests can be used as a measure of 'quality assurance' to ensure fair and appropriate grade assignment. Grades are often criticized for their subjectivity and unreliable measurement of student learning (J. Allen, 2005; Geiser & Santelices, 2007; Kohn, 1999). Much of this criticism stems from grading formulas that include student effort, participation, and work habits – factors not representative of actual student learning (J. Allen, 2005). Many schools continue, however, to assign grades for

homework, effort, and participation, despite the growing consensus amongst measurement specialists that this is not a valid form of assessment (Cross & Frary, 1999).

The example of Shanghai American School indicates that when schools invest in school-wide grading policies and discussions, grades can in fact accurately reflect the skill levels of students**.** Over the last few years, SAS separated student learning dispositions from academic grades and decided to report dispositions on a separate learner profile. Teachers are now encouraged to only include key assessments, aligned to course standards, in their grade assignments (Shanghai American School, 2011a).

This study provides evidence for successful implementation of best assessment practices and can serve as an example to other schools hoping to use grades as a valid measure of student knowledge and skills. Future research might compare these results with those of schools that do not have skills-based grading policies. Further study will help schools address issues with grading and student learning measurement.

Testing implications and future research. As discussed in the literature review, standardized tests are often criticized for their narrow assessment of student learning and focus on low-level cognitive tasks (Conley, 2003; Diamond, 2007; L. Resnick, 2010; L. Resnick & Zurawsky, 2005). The results of this study suggest, however, that academic achievement and critical thinking are so highly correlated that perhaps they cannot be approached as two separate constructs of learning. This view is supported by research on academic achievement and critical thinking (Brunt, 2005; P. Facione, 1990c; Jenkins, 1998; Vendrely, 2007). Each lends itself to the other, and teasing them out may be counterproductive.

Excellent critical thinkers may also be particularly adepts at navigating standardized tests. Beyond content knowledge, they may be able to eliminate detractor items and use their critical thinking skills to make excellent educated guesses. In many ways, test-taking skills are also critical thinking skills.

Results may also suggest that Norwest Evaluation Association, the developer of the prominent and widely used MAP test, has created a test that assesses a range of skills and can predict critical thinking ability. Mean RIT score proved to be a better predictor of critical thinking than grades. NWEA purports that their tests are aligned to state and national standards, but as of yet there is no easily accessible information on how their tests integrate thinking skills. This study provides some data on how one standardized test may integrate or predict critical thinking skills.

MAP test and critical thinking skill correlations have significant implications for schools and how they assess student learning. The MAP test is largely used to measure academic skills, but since these findings suggest it is very highly correlated with thinking skills, it may be sufficient for assuming that if students do well on it, they are also strong critical thinkers. This might ease some concerns about the costs of new test adoption; Testing can consume considerable instructional time and financial resources (Smith & Rottenberg, 1991). This study suggests that an additional test may not be necessary, and that critical thinking skills can be integrated into achievement tests. Further study and perhaps modest adjustments would be necessary, but the alignment is promising for school districts, teachers, parents, and students who would benefit from a smaller testing window that incorporates a wide range of data on thinking and learning. Schools may be

able to use the MAP test to legitimately report on student academic levels without compromising critical thinking skills as part of the overall academic profile.

Though this study sheds some light on the relationship between critical thinking and academic achievement, it remains an area for future research. It is unclear whether the MAP test measures critical thinking or that academic achievement, which it does measure, preempts critical thinking. Understanding the interaction between critical thinking and achievement can support critical thinking development for a diversity of learners.

Gender implications and future research. The results of this study support the predominant research results that find no association between critical thinking skills and gender (Ben-Chaim, et al., 2000; El Hassan & Madhum, 2007; P. Facione, 1990c; P. Facione, et al., 1995). Gender does, however, appear to influence measures of academic achievement within this sample. On average, female students receive grades nearly three points higher than their male counterparts. This may be due to differences in learning styles between males and females. Females tend to excel at building relationships through discussion, writing, and curriculum based content tasks (Halpern, 2004; Wood, 2012). These skills may be rewarded within classroom settings, but not on standardized tests.

At the same time, females tend to have lower MAP test mean scores, which is largely a result of lower performance in math. Females do slightly better in language usage, but their mean score is lower due to a five point deficit in math. Unfortunately, these findings are supported by research on women in mathematics (Catsambis, 1994;

Devine, Fawcett, Szűcs, & Dowker, 2012; Kane & Mertz, 2012; Schwartz & Hanson, 1992). The hypotheses for gender stratification are varied, but Else-Quest and colleagues' deduce from the Trends in International Mathematics and Science Study (TIMSS) and Programme for International Student Assessment (PISA) data that performance variations are largely impacted by cultural variations (Else-Quest, Hyde, & Linn, 2010). This study adds to the research with the findings that within a multicultural international context, the gender gap remains.

Gender differences in academic achievement are not the focus of this study, but the results suggest an important area for further research. Why are females earing higher grades? Why do males outperform females in math? There may be significant implications that lead to greater support for females in math studies and changes in grading policies that reward both male and female attributes.

National curriculum implications and future research. This study sought to investigate if an American school curriculum influenced critical thinking skills in students. The results found no relationship between years of schooling at SAS and critical thinking skills. Additional research would be necessary, however, to draw more concrete conclusions between the malleability of thinking skills and the impact of a Western style education. This study was limited in that it only investigated the number of years at SAS and did not take into account students who may have transferred from other international or American schools. Additionally, many new students came to SAS from Chinese international schools which are often a hybrid between Western and Chinese curriculum, pedagogy, and teacher nationality. A clear definition of Western style

education and additional data on transfer schools would be an important next step for further research.

Cross-cultural implications and future research. Data analysis reveals that native language, a proxy for culture, is not a significant independent variable for critical thinking. Students whose native language fell under the Confucian cultural cluster had the same critical thinking scores as students whose native language was classified as non-Confucian[5]. At the same time, however, when math scores are factored out, non-Confucian students do score slightly higher in critical thinking.

Results provide interesting insight into cultural differences in critical thinking as well as the role of discipline specific knowledge of critical thinking development. As discussed in the literature review, the conceptualization of critical thinking may differ across cultures. Confucian culture emphasizes judgment and self-reflection, while challenging assumptions is not part of this tradition (Biggs, 1996). Thus, Confucian students may excel in some aspects of critical thinking and not others. The results of this study support this idea. Confucian students only have equal critical thinking scores when math is taken into consideration, indicating that they are stronger, perhaps, in areas closely related to logic and mathematical reasoning. These results may also indicate that Confucian students are less skilled at transferring knowledge across disciplines. Non-Confucian high performing math students have higher overall critical thinking scores than Confucian high performing math students. According to Kuhn (1999), critical thinking can be transferred across disciplines by drawing conclusions or applying similar

[5] Native language serves as a proxy for culture in this study. Student passport, mother's passport, and father's passport were also considered but were found insignificant.

protocols. On the other hand, McPeck (1981) and to some extent cognitive psychologists (Bransford, et al. 2005), assert that critical thinking is more domain specific and is not easily transferable. The results of this study support both arguments, but add that the ease of transferability may be impacted by culture.

The overall results, which indicate that Confucian and non-Confucian students have equal scores, also have interesting implications for what is known as "the Asian Paradox." According to Biggs (1996) and Turner (2006), Asian students are often criticized for lacking reasoning skills despite the fact that they often outperform American students on international standardized tests (National Center for Education Statistics, 2013; Programme for International Student Assessment, 2012). This paradox is often rooted in an assumption that Asian students don't learn to think critically because they spend much of their class time seated in quite rows transcribing lectures by the professor (Biggs, 1996). According to this study, however, the misconception of the Asian learner is likely a result of perceived differences in learning dispositions rather than actual skill. Critical thinking dispositional differences have been recorded by Ip, et al. (2000), Ku and Ho (2010), and Tiwari, et al. (2003), finding that Chinese students express critical thinking differently. Until now, however, there has been very little research specific to critical thinking skills and Confucian culture. This study indicates the need for more research and instruction on both critical thinking skills and dispositions.

In an effort to address dispositional differences between cultures, school may decide that the California Critical Thinking Dispositions Inventory (CCTDI), also developed by Peter and Noreen Facione, is a more appropriate thinking assessment for a

particular context. With large numbers of Asian students studying in America (Altbach, 2004), it is important for students to be able to express and apply critical thinking skills in academic settings. Teaching critical thinking dispositions requires a different approach than teaching critical thinking skills. If dispositions are truly an area of concern, teachers may want to model and teach students how to formulate questions, challenge authority, and contribute to discussions (Ku & Ho, 2010). Teachers may adjust their classroom cultures to ones that promote skepticism, are flexible depending on student questioning, and are focused on open-ended group tasks (Tishman, et al., 1993). There are strategies teachers can use to support Asian students in expressing and utilizing their critical thinking skills.

Beyond critical thinking, this study also provides subject-matter implications. Consistent with many other international studies, there are strong correlations between Confucian language and math RIT scores. International research has found that students from Singapore, Taiwan, South Korea, Hong Kong, and Japan consistently score higher than American students on the Trends in International Mathematics and Science Study (TIMMS). Asian American students also outperformed their white Americans counterparts in mathematics (National Center for Education Statistics, 2013). In this study, Confucian students scored, on average, 5.6 points higher on math RIT. The cultural implications of mathematics achievement are beyond the scope of this study, but suggest an area for further research. International schools, such as SAS, present an interesting context for the study of cross-cultural education and learning. It would be

important to understand why and how some Asian students are having success in mathematics so that this success can be replicated across cultures.

Limitations

As in all research, this study is not without its limitations. The first is regarding the sample. It was conducted with a relatively small set of cross-sectional data – 8th grade students at Shanghai American School. It would be important to replicate the study at more than one school, with a wider range of student ages, and with an overall larger sample set. Secondly, this study uses two broad cultural classifications, Anglo and Confucian, based on native language. Culture is a multifaceted and complex amalgamation of an individual's education, upbringing, and heritage. This study simplified the definition of culture to make the study feasible, but it is important to note that this definition is limited. Another limitation is in the testing instrument and administration. While the CCTST-M series test is a widely used and validated measure of critical thinking, the majority of testing has been with American students in the United States. On possible explanation for cultural differences in critical thinking skills could be a test bias and this particular test's conceptualization of critical thinking. Finally, there is the issue of motivation. Though the teachers and proctor encouraged students to try their best, the results could not be tied to grades and thus some students may not have given their best effort. These limitations identify areas in which the study can be expanded or modified for future research.

Conclusion

Critical thinking skills are important life-long skills for personal freedom, financial security, and civic engagement. They are also essential to the prosperity and competitive advantage of American democracy and economic success. Critical thinking skills have been a clear educational priority since as early as Plato and Confucius, but have been more central to the educational reform agenda over the last thirty years.

Critical thinking skills are undoubtedly essential educational outcomes, but the extent of their mastery is often uncertain. Because higher order cognitive tasks, such as critical thinking, are rarely assessed, they are often deprioritized within the classroom. In an age of accountability where students and teachers are being held to higher standards for teaching and learning, educators and policy makers need to take a broader look at the measures and expectations for student achievement. Grades, standardized test scores, and cognitive skill assessments, such as critical thinking, should all be taken into account when looking at student learning outcomes. Together, these data points can provide a clearer picture of student learning – one that can help schools individualize student instruction, reflect on school assessment practices, and inform strategic planning. At the same time it would emphasize the importance of these skills that are so often a part of school missions, visions, and/or overarching values but don't get the instructional time they deserve.

This quantitative study investigates the critical thinking skill levels of eighth grade students at Shanghai American School, and the relationship between critical thinking skills and academic achievement, as measured by MAP test scores and grades.

This study also looks at potential demographic implications of gender, native language (culture), and tenure at SAS on critical thinking skills

Findings suggest that students at Shanghai American School are strong critical thinkers who score higher than the majority of test takers in grades 6-9. Critical thinking scores are strongly correlated with MAP RIT scores and grades, thus suggesting a strong relationship between critical thinking and academic achievement. Results also suggest that MAP tests and grading practices accurately reflect both the knowledge and skills, including higher order thinking skills, of students. Relationships between demographic variables and critical thinking were less strong. There was no relationship between critical thinking and gender or tenure at Shanghai American school. There was also no relationship between critical thinking and culture – except when mathematics scores were taken into consideration, non-Confucian students scored higher in critical thinking. While gender and culture do not seem to have a large impact critical thinking skills, they do appear to influence academic achievement. Females, on average, had higher grades and language usage RIT scores, but lower overall MAP test scores due to difficulty in the math section of the test. Additionally, students whose native language fell under the Confucian cultural cluster scored significantly higher in the mathematics section of the MAP test. These findings suggest relevant areas for further research. Why are females earning higher grades? Why do males outperform females in math, but score lower in language usage? Why do Confucian students find more success in math and how does that influence their critical thinking ability? Though beyond the scope of this study, this data raises some interesting questions.

In an increasingly international world of cross-cultural education, an American school in China offers a unique context for the study of critical thinking skills. Global competition as well as the desire to learn from other nations, has many pundits analyzing national school systems and seeking explanations for why, in some countries, students tend to excel (Friedman, 2006; Ripley, 2013). Critical thinking must be part of the analysis – whether as an independent assessment or embedded in current assessments. This study seeks to move forward the discussion of educational reform, assessment, and the essential role of thinking skills in the national debate.

REFERNCES

Adey, P., Csapo, B., Demetriou, A., Hautamaki, J., & Shayer, M. (2007). Can we be intelligent about intelligence? Why education needs the concept of plastic general ability. *Educational Research Review, 2*(2), 75-97.

Adler, M. (Ed.). (1984). *The paideia program*. New York: Millan Publishing Company.

Allen, D., & Bond, C. A/ v. (2001). Prepharmacy predictors of success in pharmacy school: grade point averages, pharmacy college admissions test, communication abilities, and critical thinking skills. *Pharmacotherapy: The Journal of Human Pharmacology and Drug Therapy, 21*(7), 842-849.

Allen, J. (2005). Grades as valid measures of academic achievement of classroom learning. *The Clearing House: A Journal of Educational Strategies, Issues and Ideas, 78*(5), 218-223.

Altbach, P. (2004). Higher education crosses borders: Can the United States remain the top destination for foreign students? *Change: the magazine of higher learning, 36*(2), 18-25.

Assessment, I. (2014). Sample thinking skills questions Retrieved 3/1/2014, from https://http://www.insightassessment.com/CT-Resources/node_1487

August, D. (July 31, 2012). [Email Communication].

Bailin, S. (2002). Critical thinking and science education. *Science & Education, 11*(4), 361-375.

Bailin, S., & Siegel, H. (2003). Critical thinking. *The Blackwell guide to the philosophy of education*, 181-193.

Barak, M., Ben-Chaim, D., & Zoller, U. (2006). Purposely teaching for the promotion of higher-order thinking skills: a case for critical thinking. *Research in Science Education, 13*(4), 353-369.

Ben-Chaim, D., Ron, S., & Zoller, U. (2000). The disposition of eleventh-grade science students toward critical thinking. *Journal of Science Education and Technology, 9*(2), 149-159.

Berliner, D. (2001). Learning about and learning from expert teachers. *International journal of educational research, 35*(5), 463-482.

Beyer, B. (1984). Improving thinking skills--defining the problem. *Phi Delta Kappan, 65*(7), 486-490.

Biggs, J. (1996). Western misconceptions of the Confucian-heritage learning culture. In D. Watkins & J. Biggs (Eds.), *The Chinese Learner: Cultural Psychological, and Contextual Influences*. Hong Kong: Comparative Education Research Center.

Blakey, E., & Spence, S. (1990). Developing metacognition. *ERIC Digest, ED327218*.

Bloom, B. (1956). *Taxonomy of educational objectives: The classification of educational goals* (Vol. 2). New York: Longmans.

Bok, D. (2008). *Our underachieving colleges: A candid look at how much students learn and why they should be learning more*: Princeton University Press.

Bond, M. (Ed.). (1996). *The handbook of Chinese psychology*. Hong Kong: Oxford University Press.

Brabeck, M. (1983). *Intellectual development during the college years: How strong is the longitudinal evidence?* Utara: University Utara Malaysia.

Bransford, J., Vye, N., Stevens, R., Kuhl, P., Schwartz, D., Bell, P., et al. (2005). Learning theories and education: Toward a decade of synergy. *Handbook of Educational Psychology (2nd Edition)*.

Brown, J. (1983). On teaching thinking skills in the elementary and middle school. *The Phi Delta Kappan, 64*(10), 709-714.

Brown, K. (1998). *Education, culture and critical thinking*. United Kingdom: Ashgate.

Browne, M., & Keeley, S. (1998). *Asking the right questions*: Prentice-Hall Englewood Cliffs, New Jersey.

Brunt, B. (2005). Models, measurement, and strategies in developing critical-thinking skills. *Journal of Continuing Education in Nursing, 36*(6), 255.

Bunnell, T. (2009). The international baccalaureate in the USA and the emerging 'culture war'. *Discourse: Studies in the Cultural Politics of Education, 30*(1), 61-72. doi: 10.1080/01596300802643090

Cabrera, G. (1992). A framework for evaluating the teaching of critical thinking. *Education, 113*, 59-63.

Casner-Lotto, J., & Barrington, L. (2006). *Are they really ready to work? Employers' perspectives on the basic knowledge and applied skills of new entrants to the 21st century US workforce*. Washington DC: Partnership for 21st Century Skills.

Catsambis, S. (1994). The path to math: Gender and racial-ethnic differences in mathematics participation from middle school to high school. *Sociology of Education*, 199-215.

. CCTST M- Series test manual. (2013). In M. Kelly (Ed.). San Jose, CA: California Academic Press.

Chan, N.-M., Ho, I., & Ku, K. (2011). Epistemic beliefs and critical thinking of Chinese students. *Learning and Individual Differences, 21*(1), 67-77.

Chappuis, S., & Stiggins, R. J. (2002). Classroom assessment for learning. *Educational Leadership, 60*(1), 40-44.

Chenoweth, L. (1998). Facilitating the process of critical thinking for nursing. *Nurse Education Today, 18*(4), 281-292.

Clifford, J. S., Boufal, M. M., & Kurtz, J. E. (2004). Personality traits and critical thinking skills in college students: Empirical tests of a two-factor theory. *Assessment, 11*(2), 169-176.

Cohen, M., Salas, E., & Riedel, S. (2002). Critical Thinking: Challenges, possibilities, and purpose. Ft. Leavenworth: U.S. Army Research Institute & Cognitive Technologies, Inc.

College Board. (2008). The 4th annual AP report to the nation. New York.

Conley, D. (2003). Connecting the dots: Linking high schools and postsecondary education to increase student success. *Peer Review, 5*(2), 9-12.

Conley, D. (2007). Redefining college readiness. *Educational Policy Improvement Center, 3*.

Conley, D. (2011). Building on the Common Core. *Educational Leadership, 68*(6), 16-20.

Conley, D., Lombardi, A., Seburn, M., & McGaughy, C. (2009). *Formative assessment for college readiness: Measuring skill and growth in five key cognitive strategies associated with postsecondary success*. Paper presented at the American Educational Research Association, San Diego.

Cross, L. H., & Frary, R. B. (1999). Hodgepodge grading: Endorsed by students and teachers alike. *Applied measurement in Education, 12*(1), 53-72.

Dahlin, M., Xiang, Y., Durant, S., & Cronin, J. (2010). State standards and student growth: Why state standards don't matter as much as we thought. *Northwest Evaluation Association*.

Del Bueno, D. (2005). A crisis in critical thinking. *Nursing education perspectives, 26*(5), 278-282.

Denial, A. (2008). Association of critical thinking skills with clinical performance in fourth year optometry students. *Journal of Optometry Education, 33*(3), 103-106.

Devine, A., Fawcett, K., Szűcs, D., & Dowker, A. (2012). Gender differences in mathematics anxiety and the relation to mathematics performance while controlling for test anxiety. *Behavioral and Brain Functions, 8*(1), 33.

DeVoogd, G. (2006). Question authority: Kids need to be skeptical of the curriculum. It's the only way to develop a Balanced view of the world. *School Library Journal, 52*(4), 48.

Dewey, J. (1910). *How we think*. D.C.: Heath and Company.

Dewey, J. (1966). *Democracy and education: An introduction to the philosophy of education*. New York: Free Press.

Diamond, J. B. (2007). Where the rubber meets the road: Rethinking the connection between high-stakes testing policy and classroom instruction. *Sociology of Education, 80*(4), 285-313.

El Hassan, K., & Madhum, G. (2007). Validating the Watson Glaser Critical Thinking Appraisal. *Higher Education, 54*(3), 361-383.

Else-Quest, N. M., Hyde, J. S., & Linn, M. C. (2010). Cross-national patterns of gender differences in mathematics: a meta-analysis. *Psychological bulletin, 136*(1), 103.

Ennis, R. (1980). A conception of rational thinking. In J. R. Coombs (Ed.), *Philosophy of Education* Normal: Philosophy of Education Society.

Ennis, R. (1985). A logical basis for measuring critical thinking skills. *Educational Leadership, 43*(2), 44-48.

Ennis, R. (1993). Critical thinking assessment. *Theory Into Practice, 32*(3), 179-186. doi: 10.1080/00405849309543594

Erwin, T. D. (2000). The NPEC sourcebook on assessment, volume 1: Definitions and assessment methods for critical thinking, problem solving and writing [and] Volume 2: Selected institutions utilizing assessment results: National Postsecondary Education Cooperative.

Facione, N., & Facione, P. (1996). Assessment design issues for evaluating critical thinking in nursing. *Holistic Nursing Practice, 10*(3), 41-53.

Facione, N., & Facione, P. (1997). *Critical thinking assessment in nursing education programs: An aggregate data-analysis*. Millbrae: California Academic Press.

Facione, P. (1990a). The California critical thinking skills test--college level. Technical report #1. Experimental validation and content validity. Milbrae.

Facione, P. (1990b). The California critical thinking skills test--college level. Technical report #2: Factors predictive of CT skills. Milbrae.

Facione, P. (1990c). The California critical thinking skills test--college level. Technical report No. 3. Gender, ethnicity, major, CT self-esteem, and the CCTST. Milbrae.

Facione, P. (1990d). Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction. Research findings and recommendations. Newark: American Philosophical Association.

Facione, P. (2000). The disposition toward critical thinking: Its character, measurement, and relationship to critical thinking skill. [Journal Article]. *Informal Logic, 20*(1), 61-84.

Facione, P. (2011). Critical Thinking: What it is and why it counts. Millbrae: Insight Assessment.

Facione, P., & Facione, N. (2007). *Thinking and reasoning in human decision making: The method of argument and heuristic analysis*. Milbrae: Insight Assessment.

Facione, P., Gittens, C. A., Facione, N., & Winterhalter, K. (2011). CCTST M-series test manual. Millbrae: The California Academic Press.

Facione, P., Sanchez, C., Facione, N., & Gainen, J. (1995). The disposition toward critical thinking. *The Journal of General Education, 44*(1), 1-25.

Facione, P., (1990). Critical Thinking: A Statement of Expert Consensus for Purposes of Educational Assessment and Instruction. Research Findings and Recommendations.

Fisher, A. (2001). *Critical thinking : An introduction*. Cambridge: Cambridge University Press.

Florida, R. (2006). The flight of the creative class: the new global competition for talent. *Liberal Education, 92*(3), 22.

Forehand, M. (2010). Bloom,Äôs taxonomy. *Retrieved July, 28*, 2010.

Freire, P. (1974). *Pedagogy of the oppressed*. New York: Seabury Press.

Friedman, T. L. (2006). *The world is flat [updated and expanded]: A brief history of the twenty-first century*: Farrar, Straus and Giroux.

Gaff, J. (2004). What is a generally educated person? *Peer Review*, 4-7.

Garett, K., & Wulf, K. (1978). Relationship of a measure of critical thinking ability to personality variables and to indicators of academic achievement. *Educational & Psychological Measurement, 38*, 1181-1187.

Geiser, S., & Santelices, M. V. (2007). Validity of high-school grades in predicting student success beyond the freshman year: High-school record vs. standardized tests as indicators of four-year college outcomes.

Gelder, T. V. (2005). Teaching critical thinking: Some lessons from cognitive science. *College teaching, 53*(1), 41-48.

Giancarlo, C., & Facione, P. (2001). A look across four years at the disposition toward critical thinking among undergraduate students *The Journal of General Education, 50*(1), 29-55.

Giddens, J., & Gloeckner, G. (2005). The relationship of critical thinking performance on the NCLEX-RN. *Journal of Nursing Education, 44*(2), 85-89.

Girot, E. A. (1995). Preparing the practitioner for advanced academic study: The development of critical thinking. *Journal of Advanced Nursing, 21*(2), 387-394. doi: 10.1111/1365-2648.ep8558175

Greenhill, V. (2010). 21st Century Knowledge and Skills in Educator Preparation. *Partnership for 21st Century Skills*.

Gutmann, A. (1995). Civic education and social diversity. *Ethics*, 557-579.

Gutmann, A., & Thompson, D. (2004). *Why deliberative democracy?* : Princeton University Press.

Halpern, D. F. (1998). Teaching critical thinking for transfer across domains: Disposition, skills, structure training, and metacognitive monitoring. *American Psychologist, 53*(4), 449.

Halpern, D. F. (2001). Assessing the effectiveness of critical thinking instruction. *The Journal of General Education, 50*(4), 270-286.

Halpern, D. F. (2004). A cognitive-process taxonomy for sex differences in cognitive abilities. *Current directions in psychological science, 13*(4), 135-139.

Healy, J. M. (1990). *Endangered minds: Why children don't think and what we can do about it*: Simon and Schuster.

Heise, M. (1994). Goals 2000 Educate America Act: The federalization and Legislation of educational policy. *Fordham L. Rev., 63*, 345.

Helms, R. (2008). Transnational education in China: Challenges, critical issues, and strategies for success. *International Higher Education, 53*.

Hess, K. K., Carlock, D., Jones, B., & Walkup, J. R. (2009). What exactly do "fewer, clearer, and higher standards" really look like in the classroom? Using a cognitive rigor matrix to analyze curriculum, plan lessons, and implement assessments. *Hess' Local Assessment Toolkit: Exploring Cognitive Rigor*.

Hofer, B. K., & Pintrich, P. R. (1997). The development of epistemological theories: Beliefs about knowledge and knowing and their relation to learning. *Review of educational research, 67*(1), 88-140.

Hofstede, G. (1983). The cultural relativity of organizational practices and theories. *Journal of international business studies, 14*(2), 75-89.

Hofstede, G., Hofstede, G. J., & Minkov, M. (2010). *Cultures and organizations: Software of the mind 3rd edition*. New York: McGraw-Hill.

Hopkins, C., Raymond, M. A., & Carlson, L. (2011). Educating students to give them a sustainable competitive advantage. *Journal of Marketing Education, 33*(3), 337-347. doi: 10.1177/0273475311420241

Huitt, W. (1995). Success in the information age: A paradigm shift. *Valdosta State University, 6*.

Huitt, W. (1998). Critical thinking: An overview. *Educational Psychology Interactive*.

International Baccalaureate North America. (2007). The IB diploma programme statistical bulletin Retrieved June 26, 2008, from http://www.ibo.org.

Ip, W., Lee, D., Lee, I., Chau, J., Wootton, Y., & Chang, A. (2000). Disposition towards critical thinking: A study of Chinese undergraduate nursing students. *Journal of Advanced Nursing, 32*(1), 84-90. doi: 10.1046/j.1365-2648.2000.01417.x

Irwin, T. (1999). Aristotle: Nicomachean Ethics. *Indianapolis, IN: Hackett*.

Jenkins, E. (1998). The significant role of critical thinking in predicting auditing students' performance. *Journal of Education for Business, 73*(5), 274-279.

Kane, J. M., & Mertz, J. E. (2012). Debunking myths about gender and mathematics performance. *Notices of the AMS, 59*(1), 10-21.

Kay, K. (2009). Middle schools preparing young people for 21st century life and work. *Middle School Journal, 40*(5), 41-45.

Kennedy, M., Fisher, M. B., & Ennis, R. H. (1991). Critical thinking: Literature review and needed research. *Educational values and cognitive instruction: Implications for reform*, 11-40.

King, P., & Kitchener, K. (1994). *Developing reflective judgment: Understanding and promoting intellectual growth and critical thinking in adolescents and adults* San Francisco: Jossey-Bass.

King, P., Wood, P., & Mines, R. (1990). Critical thinking among college and graduate students. *Review of Higher Education, 13*(2), 167-186.

Kohn, A. (1999). From degrading to de-grading. *High School Magazine, 6*(5), 38-43.

Krathwohl, D. R. (2002). A revision of Bloom's taxonomy: An overview. *Theory Into Practice, 41*(4), 212-218.

Ku, K., & Ho, I. (2010). Dispositional factors predicting Chinese students' critical thinking performance. *Personality & Individual Differences, 48*(1), 54-58. doi: 10.1016/j.paid.2009.08.015

Kuhn, D. (1999). A developmental model of critical thinking. *Educational Researcher, 28*(2), 16-46.

Kuhn, D. (2000). Metacognitive development. *Current directions in psychological science, 9*(5), 178-181.

Kurfiss, J. G. (1988). Critical thinking: Theory, research, practice, and possibilities. Washington D.C.: Association for the Study of Higher Education.

Kwon, N. (2008). A mixed-methods investigation of the relationship between critical thinking and library anxiety among undergraduate students in their information search process. *College & Research Libraries, 69*(2), 117-131.

Lai, E. (2011). Critical thinking: A literature review. *Pearson research report*.

Leach, B. (2011). *Critical thinking skills as related to university students' gender and academic discipline.* Doctorate of Education, East Tennessee State University.

Lee, W.O. (1996). The Cultural Context for Chinese Learners: Conceptions of Learning in the Confucian Tradition. In D. Watkins & J. Biggs (Eds.), *The Chinese Learner: Cultural, Psychological, and Contextual Influences*. Hong Kong: Comparative Education Research Center.

Linn, R. L. (2000). Assessments and accountability. *Educational Researcher, 29*(2), 4-16.

Mansfield, H. (2001). Grade inflation: It's time to face the facts. *Chronicle of Higher Education, 47*(30), B24.

. MAP overview. (2012) Retrieved August 15, 2012, 2012, from http://www.nwea.org/

McCollister, K., & Sayler, M. F. (2010). Lift the ceiling. *Gifted Child Today, 33*, 41-47.

McElroy, S. (October 3, 2012). [email communication].

McLester, S., & McIntire, T. (2006). The workforce readiness crisis. *Technology & Learning, 27*(4), 22-29.

McMullan, M., Endacott, R., Gray, M. A., Jasper, M., Miller, C. M. L., Scholes, J., et al. (2003). Portfolios and assessment of competence: A review of the literature. *Journal of Advanced Nursing, 41*(3), 283-294.

McPeck, J. (1981). *Crtical thinking and education*. New York: Martin's Press.

Milton, G., & Harvey, J. (1983). A Nation at Risk: The Report of the National Commission on Excellence in Education. *The Phi Delta Kappan, 65*(1), 14-18.

More Korean children sent to study abroad alone. (2009). *The Chosun Libo*. Retrieved from http://english.chosun.com/site/data/html_dir/2006/03/06/2006030661022.html

National Center for Education Statistics. (2013). Highlights from TIMSS 2011 (pp. 56): U.S. Department of Education.

National Commission on Excellence in Education. (1983). A nation at risk: The imperative for educational reform. *The Elementary School Journal*, 113-130.

National Research Council. (2004). *How people learn: Brain, mind, experience, and school*. Washington D.C.: National Academy Press.

NCEA. (2010). The advanced placement program benefits mainly well-prepared students who pass AP exams: Issues in college readiness: National Center for Educational Achievement.

Neisser, U., Boodoo, G., Bouchard, T., Boykin, W., Brody, N., Ceci, S., et al. (1996). Intelligence: Knowns and unknowns. *American Psychologist, 51*(2), 77-101.

Nelsen, P., & Seaman, J. (2011). Deweyan tools for inquiry and the epistemological context of critical pedagogy. *Educational Studies: Journal of the American Educational Studies Association, 47*(6), 561-582.

Norris, S. (1985). Synthesis of research on critical thinking. [Article]. *Educational Leadership, 42*(8), 40.

Norris, S. (1988). Controlling for background beliefs when developing multiple-choice critical thinking tests. [Article]. *Educational Measurement: Issues & Practice, 7*, 5-11. doi: 10.1111/j.1745-3992.1988.tb00437.x

Northwest Evaluation Association. (2003). Technical manual for the NWEA measures of academic progress and achievement level tests. *Computer software manual*.

Northwest Evaluation Association. (2011). A study of the alignment of the RIT scales of NWEA's MAP assessments with the college readiness benchmarks of EXPLORE, PLAN, and ACT *College Readiness Linking Study*.

O'Reilly, K. (2010). The 'Apology': An Introduction to the dynamics of the moral life. [Journal Article]. *Journal of Philosophy and Religion, 11*(2), 35-52.

Olson, L. (2003). Standards and tests: Keeping them aligned. *Research points: Essential information for education policy, 1*(1).

Partnership for 21st Century Skills, & American Association of Colleges of Teacher Education. (2010). 21st century knowledge and skills in educator preparation. In V. Greenhill (Ed.): ERIC.

Paul, R. (1985a). Bloom's taxonomy and critical thinking instruction. *Educational Leadership, 42*(8), 36-39.

Paul, R. (1985b). McPeck's mistakes. *Informal Logic, 7*, 35-43.

Paul, R. (1992). Critical thinking: What, why, and how. *New Directions for Community Colleges*(77), 3-24.

Paul, R., & Binker, A. (1990). *Critical thinking: What every person needs to survive in a rapidly changing world*. Sonoma: Univ, Rohnert Park C. A. Center for Critical Thinking.

Paul, R., Elder, L., & Bartell, T. (1997). California teacher preparation for instruction in critical thinking: Research findings and policy recommendations: California Commission on Teacher Credentialing.

Programme for International Student Assessment. (2012). PISA 2012 results in focus (pp. 32): OECD.

Purdy, L. M. (1992). *In their best interest?: the case against equal rights for children*: Cornell University Press.

Resnick, D., & Peterson, N. (1991). Evaluating progress toward goal five: A report to the National Center for Education Statistics.

Resnick, L. (2010). Nested learning systems for the thinking curriculum. [Article]. *Educational Researcher, 39*(3), 183-197. doi: 10.3102/0013189x10364671

Resnick, L., & Zurawsky, C. (2005). Getting back on course. *American Educator, 29*(1), 9-19.

Richmond, B. (1993). Systems thinking: critical thinking skills for the 1990s and beyond. *System Dynamics Review, 9*(2), 113-133.

Ripley, A. (2013). *The smartest kids in the world: And how they got that way*: Simon and Schuster.

Rothman, R., Slattery, J., Vranek, J., & Resnick, L. (2002). *Benchmarking and alignment of standards and testing*: Center for the Study of Evaluation, National Center for Research on Evaluation, Standards, and Student Testing, Graduate School of Education & Information Studies, University of California, Los Angeles.

Rothschild, E. (1999). Four decades of the advanced placement program. *History Teacher, 32*(2), 175-206.

Santoli, S. P. (2002). Is there an Advanced Placement advantage? *American Secondary Education, 30*(3), 23-35.

Sawaya, V. C. (2012). The truth of the matter: A defense of critical thinking as the principle aim of education. [Journal Article]. *Stance: An International Undergraduate Philosophy Journal, 5*, 75-84.

Schwartz, W., & Hanson, K. (1992). *Equal mathematics education for female students*: ERIC Clearinghouse on Urban Education.

Scott, J. N., & Markert, R. J. (1994). Relationship between critical thinking skills and success in preclinical courses. *Academic Medicine*.

Shanghai American School. (2011a). *How we are approaching learning at Puxi MS*. Shanghai American School. Shanghai.

Shanghai American School. (2011b). *SAS assessment philosophies, practices, and procedures manual*. Shanghai American School. Shanghai.

Shanghai American School. (2012). Shanghai American School core values and mission Retrieved August 7, 2012, 2012, from http://www.saschina.org/?page=corevalues_mission

. Shanghai American School middle school program. (2012) Retrieved August 15, 2012, 2012, from http://www.saschina.org/?page=Middle_School_Progra

Sharifi, M., McCombs, G. B., Fraser, L. L., & McCabe, R. K. (2009). Structuring a competency-based accounting communication course at the graduate level. [Article]. *Business Communication Quarterly, 72*(2), 177-199.

Siegel, H. (1980). Critical thinking as an educational ideal. [Article]. *Educational Forum, 45*, 7-23.

Siegel, H. (1988). *Educating reason: Rationality, critical thinking, and education*. New York: Routledge.

Silva, E. (2009). Measuring Skills for 21st-Century Learning. *The Phi Delta Kappan, 90*(9), 630-634.

Simon, A. W., L. (1974). The performance on the Watson Glaser Critical Thinking Appraisal of university students classified according to sex, type of course pursued, and personality score category. *Educational and Psychological Measurement, 34*(4), 957-960.

Smith, M. L., & Rottenberg, C. (1991). Unintended consequences of external testing in elementary schools. *Educational measurement: Issues and practice, 10*(4), 7-11.

Srinivasan, S., & Cooks, S. (2005). *Does gender influence critical thinking attitudes?* Paper presented at the Proceeding of Society for Information Technology & Teacher Education International Conference 2005, Cesapeake, V.A.

Sternberg, R. J. (1986). Critical thinking: Its nature, measurement, and improvement.

Steward, R., & Al-Abdulla, Y. (1989). An examination of the relationship between critical thinking and academic success on a university campus. Washington D.C.: U.S. Department of Education.

Strahan, D. B. (1986). Guided thinking: A research-based approach to effective middle grades instruction. *The Clearing House, 60*(4), 149-155.

Sullivan-Mann, J., Perron, C., & Fellner, A. (2009). The effects of simulation on nursing students' critical thinking scores: a quantitative study. *Journal of Nurse Midwifery & Women's Health, 9*(2), 111-116.

Sun, W., & Boncella, R. J. (2007). Transnational higher education across borders: Issues effecting joint degree programs among US and Chinese schools. *Issues in Information Systems, 8*(1), 65-70.

Taylor, M. L., & Porath, M. (2006). Reflections on the International Baccalaureate Program: Graduates' perspectives. [Article]. *Journal of Secondary Gifted Education, 17*(3), 149-158.

Ten Dam, G., & Volman, M. (2004). Critical thinking as a citizenship competence: Teaching strategies. *Learning and Instruction, 14*, 359-379.

Thayer-Bacon, B. J. (1993). Caring and its relationship to critical thinking. [Article]. *Educational Theory, 43*, 323-340. doi: 10.1111/j.1741-5446.1993.00323.x

Thayer-Bacon, B. J. (2000). *Transforming critical thinking: thinking constructively*. New York: Teachers College Press.

This we believe: Successful schools for young adolescents. (2003). Westerville: National Middle School Association.

Tishman, S., Jay, E., & Perkins, D. N. (1993). Teaching thinking dispositions: From transmission to enculturation. *Theory Into Practice, 32*(3), 147-153.

Tiwari, A., Avery, A., & Lai, P. (2003). Critical thinking disposition of Hong Kong Chinese and Australian nursing students. [Article]. *Journal of Advanced Nursing, 44*(3), 298-307. doi: 10.1046/j.1365-2648.2003.02805.x

Tombari, M. L., & Borich, G. D. (1999). *Authentic assessment in the classroom: Applications and practice*. New Jersey: Prentice-Hall.

Turner, Y. (2006). Students from mainland China and critical thinking in postgraduate Business and Management degrees: teasing out tensions of culture, style and substance. *International Journal of Management Education, 5*(1), 3-11.

United States Congress. (1994). *Goals 2000: Educate America Act*. Washington D.C: Retrieved from http://www2.ed.gov/legislation/GOALS2000/TheAct/index.html.

Valiga, M. J. (1990). *Assessing the reliability of survey instruments.* Ed.D., University of Iowa. Retrieved from http://search.ebscohost.com/login.aspx?direct=true&db=eric&AN=ED232579&site=ehost-live Available from EBSCOhost eric database.

Van Horn, R. (2003). Computer adaptive tests and computer-based tests. *The Phi Delta Kappan, 84*(8), 567-631.

Vendrely, A. (2007). An Investigation of the Relationships Among Academic Performance, Critical Performance, Critical Thinking, and Success on the Physical Therapy Licensure Examination. *Journal of Allied Healthy, 36*(2), 108-123.

Vobejda, B. (1988, February 26, 1988). Students Reported Deficient In Critical-Thinking Skills, *Washington Post*. Retrieved from http://www.highbeam.com/doc/1P2-1241959.htmlhttp://www.highbeam.com/doc/1P2-1241959.html

Volet, S., & Renshaw, P. (1996). Chinese students at an Australian university:adaptability and continuity. In D. Watkins & J. Biggs (Eds.), *The Chinese Learner: Cultural, Psychological, and Contextual Influences*. Hong Kong: Comparative Education Research Centre.

Wagner, T. (2008). Rigor redefined. *Educational Leadership, 66*(2), 20-24.

Wagner, T. (2010). *The global achievement gap: Why even our best schools don't teach the new survival skills our children need--and what we can do about it*: Basic Books.

Wallis, C., & Steptoe, S. (2006). How to bring our schools out of the 20th century. *Time magazine, 168*(25), 50-56.

Walsh, D., & Paul, R. W. (1986). *The Goal of Critical Thinking: from Educational Ideal to Educational Reality*. Washington D. C: American Federation of Teachers.

Wang, S., McCall, M., Jiao, H., & Harris, G. (2012). Construct Validity and Measurement Invariance of Computerized Adaptive Testing: Application to Measures of Academic Progress (MAP) Using Confirmatory Factor Analysis. *Online Submission*.

Way, W. D. (2006). Online Testing Research: Informing and Guiding Transitions to Computerized Assessment.

Weinstein, M. (1991). Critical thinking and education for democracy: Montclair State College Institution for Critical Thinking.

Western Interstate Commission for Higher Education. (2006). Accelerated learning options: Moving the needle on access and success. Retrieved from http://www.wiche.edu/Policy/Accelerated_Learning/

White, B. Y., & Frederiksen, J. R. (1998). Inquiry, Modeling, and Metacognition: Making Science Accessible to All Students. *Cognition and Instruction, 16*(1), 3-118.

Wiggins, G. (1989). A true test: Toward more authentic and equitable assessment. *Phi Delta Kappan, 70*(9), 703-713.

Wiggins, G., & McTighe, J. (2005). *Understanding by design*. New Jersey: Pearson.

Williams, K., Glasnap, D., Tilliss, T., Osborn, J., Wilkins, K., Mitchell, S., et al. (2003). Predictive validity of critical thinking skills for initial clinical dental hygiene performance. *Journal of Dental Education, 67*(11), 1180-1192.

Williams, R. (2003). Critical thinking as a predictor and outcome measure in a large undergraduate educational psychology course.

Williams, R., Oliver, R., Allin, J., Winn, B., & Booher, C. (2003). Psychological critical thinking as a course predictor and outcome variable. *Teaching of Psychology, 30*(3), 220.

Williams, R., & Stockdale, S. (2003). High-Perfroming Students With Low Critical Thinking Skills. *The Journal of General Education, 52*(3), 200-226.

Willingham, D. T. (2008). Critical Thinking: Why Is It So Hard to Teach? *Arts Education Policy Review, 109*(4), 21-32.

Wood, J. T. (2012). *Gendered lives*: Wadsworth Publishing Company.

Woodruff, D. J., & Ziomek, R. L. (2004). *High school grade inflation from 1991 to 2003*: ACT.

Xiang, Y., & Hauser, C. (2010). School Conditional Growth Model: How to Make an" Apples to Apples" Comparison Possible? *Northwest Evaluation Association*.

Yang, A., Ya-Ting, C., Newby, T., & Bill, R. (2008). Facilitating interactions through structured web-based bulletin boards: A quasi-experimental study on promoting learners' critical thinking skills. *Computers and Education, 50*(4), 1572-1585.

Zhang, L.-F. (2003). Contributions of thinking styles to critical thinking dispositions. [Article]. *Journal of Psychology, 137*(6), 517-543.

Ziomek, R. L., & Svec, J. C. (1995). High School Grades and Achievement: Evidence of Grade Inflation. ACT Research Report Series 95-3.

APPENDICES

Appendix A: Sample Thinking Skills Questions,
Instructions: Form a reflective and reasoned judgment with regard to which choice is the best from among those offered.
Sample Test Items 1-3

For Sample Items 1, 2 and 3 Please consider this information: A scientific study compared two matched groups of college women. The women in both groups were presented with information about the benefits of a healthy diet and regular exercise. The women in one group were paired up with one another and encouraged to work as two-person teams to help each other stick with the recommended healthy regimen of smart eating and regular vigorous exercise. The women in the other group were encouraged to use the same recommended regimen, but they were also advised to work at it individually, rather than with a partner or teammate. After 50 days the physical health and the well-being of all the women in both groups were evaluated. On average the women in the first group (with teammates) showed a 26 point improvement in measures of cardiopulmonary capacity, body strength, body fat reduction, and sense of well-being. On average the women in the other group (encouraged to work as individuals) showed a 17 point improvement on those same measures. Using statistical analyses the researchers determined that the probability that a difference of this size had occurred by chance was less than one in 1000.

Sample Item # 1.

If true, these research findings would tend to support which of the following assertions?
A = A college woman cannot achieve optimal health functioning without a teammate.
B = Universities should require all students living in campus residence halls to participate in a health regime of smart eating and regular vigorous exercise.

C = A healthy diet will cause one to have better mental health and physical strength.
D = This research study was funded by a corporation that makes exercise apparel.
E = A regimen of smart eating and regular exercise is related to better health.

Sample Item # 2.

If the information given in the case above were true, which of the following hypotheses would not need to be ruled out in order to confidently claim that for the majority of young adults a regimen of smart eating and regular vigorous exercise will result in significant improvements in one's overall health.
A = This study was about women, the findings cannot be generalized to include men.
B = Since the study began to solicit willing participants before the Research Ethics Review Committee of the college gave the research project its formal approval to gather data, the findings are invalid.
C = Some women in the study over-reported their compliance with the eating and exercise regimen, which led the researchers to underestimate the full impact of the regimen.
D = Since many of those studied described themselves as overweight or out of shape when the study began, a similar regimen will not benefit people who are healthier to start with.
E = The performance tests used to evaluate the health and well-being of females may not be appropriate for evaluating the health and well-being of males.

Sample Item # 3.

Consider the claim, "Working with a teammate or partners on a health regimen is better than working individually." Which of the following additional pieces of information would not weaken that claim?
A = Most of the women in the group that was encouraged to work individually actually worked with friends and partners who were not part of the study.
B = Most of the pairings and teams created in the first group (with teammates) fell apart after a few days and the women in that group actually worked individually.
C = There was something about the women in the first group (with teammates) that the researchers overlooked, thus invalidating the intended matching of the two groups.
D = Men are more likely to work alone, so any recommendation that men find a teammate or partner to support them in sticking with the regimen will be ignored.
E = The study was undertaken when there were no exams or major projects due, thus the results about working with a teammate do not apply to more stressful times of the year.

(Assessment, 2014)

Appendix B: Demographic Questions

1. What is your gender?
male female

2. How long have you been at SAS?
<1 year 1-2 years 3-4 years 5+ years

3. What school, if any, did you attend prior to SAS? Please indicate the name and country of the school.

4. What is the primary language spoken in your home?
English Mandarin/Cantonese Korean Japanese Hindi Spanish German French Swedish Other Asian Other European Other(please specify) _____

5. What country issued YOUR passport?
United States Canada South Korea PRC Hong Kong Taiwan Singapore Japan India New Zealand Australia Other-Southeast Asia Other-South America Other-Europe I don't know Other: _____

6. What country issued your MOTHER's passport?
United States Canada South Korea PRC Hong Kong Taiwan Singapore Japan India New Zealand Australia Other-Southeast Asia Other-South America Other-Europe I don't know Other: _____

7. What country issued your FATHER's passport?
United States Canada South Korea PRC Hong Kong Taiwan Singapore Japan India New Zealand Australia Other-Southeast Asia Other-South America Other-Europe I don't know Other: _____

8**. How long have you lived in China?**
0-1 year 2-4 years 5-8 years 9+ years

9. How many different countries have you lived for a year or more?
1 2 3 4 5 6+

Appendix C: Correlation Matrixes

Table 1: ***Regression of Critical Thinking on Demographic Variables and Grades (N=282)***

		GradeScore
Pearson Correlation	CT_overall	.583
	GenderCT	.284
	language_hindi_nonConfu_1	.049
	mean_num_years	.019
	GradeScore	1.000
Sig. (1-tailed)	CT_overall	.000
	GenderCT	.000
	language_hindi_nonConfu_1	.205
	mean_num_years	.374
	GradeScore	.
N	CT_overall	282
	GenderCT	282
	language_hindi_nonConfu_1	282
	mean_num_years	282
	GradeScore	282

** Correlation is significant at the 0.01 level (2-tailed).

Table 2: ***Regression of Critical Thinking on Demographic Variables, Grades, and Mean RIT Score (N=274)***

		CT_overall	GenderCT	language_hindi_nonConfu_1	mean_num_years
Pearson Correlation	CT_overall	1.000	.064	.014	.004
	GenderCT	.064	1.000	-.088	-.013
	language_hindi_nonConfu_1	.014	-.088	1.000	.406
	mean_num_years	.004	-.013	.406	1.000
	GradeScore	.582	.285	.047	.021
	average RIT	.721	.059	.065	.012
Sig. (1-tailed)	CT_overall	.	.145	.406	.471
	GenderCT	.145	.	.074	.413
	language_hindi_nonConfu_1	.406	.074	.	.000
	mean_num_years	.471	.413	.000	.
	GradeScore	.000	.000	.221	.365
	average RIT	.000	.163	.140	.420
N	CT_overall	274	274	274	274
	GenderCT	274	274	274	274
	language_hindi_nonConfu_1	274	274	274	274
	mean_num_years	274	274	274	274
	GradeScore	274	274	274	274
	average RIT	274	274	274	274

** Correlation is significant at the 0.01 level (2-tailed).

Table 3: ***Regression of Critical Thinking on Demographic Variables, Grades, and Math RIT Score (N=275)***

		CT_overall	GenderCT	language_hindi_nonConfu_1	mean_num_years
Pearson Correlation	CT_overall	1.000	.065	.013	.005
	GenderCT	.065	1.000	-.091	-.009
	language_hindi_nonConfu_1	.013	-.091	1.000	.401
	mean_num_years	.005	-.009	.401	1.000
	GradeScore	.582	.284	.048	.020
	Math.TestRITScore	.627	-.071	.247	.114
Sig. (1-tailed)	CT_overall	.	.141	.412	.464
	GenderCT	.141	.	.066	.438
	language_hindi_nonConfu_1	.412	.066	.	.000
	mean_num_years	.464	.438	.000	.
	GradeScore	.000	.000	.216	.371
	Math.TestRITScore	.000	.121	.000	.029
N	CT_overall	275	275	275	275
	GenderCT	275	275	275	275
	language_hindi_nonConfu_1	275	275	275	275
	mean_num_years	275	275	275	275
	GradeScore	275	275	275	275
	Math.TestRITScore	275	275	275	275

** Correlation is significant at the 0.01 level (2-tailed).

Appendix D: Regression of Critical Thinking on Demographic Variables, Grades, and Language Usage RIT Score (N=276*)*

Predictors		*Standardized Coefficients-Beta*	*t*	*Sig.*	R^2
1	(Constant)		80.814	.000	
	Gender	..070	1.151	.251	
	Language	.025	.380	.704	
	Tenure at SAS	-.007	-.102	.919	
					.005
	F = .468***				
2	(Constant)		-3.410	.001	
	Gender	-.113	-2.395	.017	
	Language	.003	.055	.956	
	Tenure at SAS	.009	.183	.855	
	Grades	.309	4.881	.000	
	Language Usage RIT	.447	7.239	.000	
					.457
	F change = 112.275 ***				

Appendix E: Regression of Critical Thinking on Demographic Variables, Grades, and Reading RIT Score (N=274)

Predictors		*Standardized Coefficients-Beta*	*t*	*Sig.*	R^2
1	(Constant)		80.166	.000	
	Gender	..066	1.084	.279	
	Language	.022	.325	.745	
	Tenure at SAS	-.004	-.053	.958	
					.005
	F = .411***				
2	(Constant)		-2.802	.005	
	Gender	-070	-1.479	.140	
	Language	.024	.484	.629	
	Tenure at SAS	.031	.631	.529	
	Grades	.356	6.135	.000	
	Reading RIT	.425	7.579	.000	
					.465
	F change = 115.478 ***				

Appendix F: Tenure on Critical Thinking, Grades, and RIT Scores

Dependent		*Mean Square*	*F*	*Sig.*
Critical Thinking	Between Groups	15.248	.363	.780
	Within Groups	41.996		
Grades	Between Groups	26.751	.842	.472
	Within Groups	31.772		
Mean RIT	Between Groups	61.385	.758	.519
	Within Groups	81.004		
Reading RIT	Between Groups	103.344	1.095	.352
	Within Groups	94.393		
Math RIT	Between Groups	413.893	2.518	.058
	Within Groups	164.360		
Language RIT	Between Groups	73.991	1.035	.377
	Within Groups	71.485p		