
2015

Assessing Numeracy in the Upper Elementary and Middle School Years

Carol Ann Gittens
Santa Clara University, cgittens@scu.edu

Follow this and additional works at: <https://digitalcommons.usf.edu/numeracy>

 Part of the [Educational Assessment, Evaluation, and Research Commons](#), and the [Science and Mathematics Education Commons](#)

Recommended Citation

Gittens, Carol Ann. "Assessing Numeracy in the Upper Elementary and Middle School Years." *Numeracy* 8, Iss. 1 (2015): Article 3. DOI: <http://dx.doi.org/10.5038/1936-4660.8.1.3>

Authors retain copyright of their material under a [Creative Commons Non-Commercial Attribution 4.0 License](#).

Assessing Numeracy in the Upper Elementary and Middle School Years

Abstract

Numeracy is the ability or tendency to reason critically about quantitative information. The preponderance of published research on numeracy examines this construct among either pre-K or early elementary samples, students with developmental challenges, or is focused on post-secondary and adult cohorts. The numeracy skills of upper-elementary and middle school students is less documented and understood, most notably because of the lack of valid instruments that are developmentally appropriate for the age range. A numeracy scale for use among upper-elementary and middle school students is introduced in this paper. Scale validation was performed using a gender-balanced, racially / ethnically diverse sample of 3rd through 8th grade students (N=197) from a private, Catholic K-8 school in Santa Clara County, California. Construct validity is supported by strong, positive correlations with the three subscales of the Cognitive Abilities Test (CogAT) as well as the Iowa Test of Basic Skills Mathematics test, a standardized academic achievement domain assessment. A preliminary exploration of the critical-thinking dispositional correlates of numeracy suggests a positive relationship with students' self-reported creative problem solving, diligence, systematicity, and fair-mindedness.

Keywords

numeracy, quantitative literacy, assessment, critical thinking, dispositions, adolescents

Creative Commons License



This work is licensed under a [Creative Commons Attribution-Noncommercial 4.0 License](https://creativecommons.org/licenses/by-nc/4.0/)

Cover Page Footnote

Carol Ann Gittens, a social and personality psychologist, is an associate professor of liberal studies at Santa Clara University (CA). She currently serves as associate dean in the College of Arts and Sciences. She was the founding director of the Santa Clara University Office of Assessment. She is a member of the American Educational Research Association, the American Psychological Association, and the National Council of Teachers of Mathematics. She is a senior research associate with Insight Assessment, LLC.

Introduction

There is a societal need for a numerically literate people in today's Information Age. Children, adolescents, and adults alike need to be able to think critically about the mathematical and numerical information that surrounds us in the media, on the Internet, in schools and workplaces, and in society at large. Steen (1997; 1999; 2000a) and others (Hauser et al. 2005; Pugalee et al. 2008; Wilkins 2000) have argued that the construct of numeracy rivals reading literacy and language fluency in its importance as a vital life skill and habit of mind. Numerically literate individuals understand the social and pragmatic function of mathematics and have the ability to reason about quantitative information. Numeracy, or its empirical synonyms, mathematical reasoning or quantitative literacy (Steen, 1990, 1999), is defined here as the ability or tendency to critically reason about quantitative information. Numeracy is needed to understand our data-driven world, to be successful in the workplace, to achieve academically, to be engaged citizens, and to make thoughtful and well-supported fiscal, social, educational, and health-oriented decisions in our personal lives (Rivera-Batiz 1992; Steen 2000b; Wilkins 2000; Wiest et al. 2007; Root 2009; Gittens and Facione 2014).

Despite the provocative socio-political rationale for orienting our educational systems around the development of critical literacies including numeracy, surprisingly little research has been conducted to date on the developmental window where children are transitioning from concrete operational thought to formal operations including clarification, inference and application of representations and abstractions (Gray et al. 1999; Koedinger et al. 1999; Ojose 2008). A search of recent peer-reviewed domestic (United States) and international publications reveals that the vast majority of studies of numeracy / quantitative literacy focuses on emergent numeracy and literacy skills among pre- and Kinder children (e.g., Methe et al. 2008; Purpura and Lonigan 2013; Skwarchuk et al. 2014), numeracy among children with disabilities (e.g., Foegen 2008; Kritzer 2012; Chan et al. 2013; Mononen et al. 2014) or the application of numeracy skills to college or adult contexts (e.g., Ward et al. 2011; van de Mortel et al. 2014; Yee and Simon 2014). One also sees an emerging literature focusing on integrating numeracy across the secondary curriculum (Crowe 2010; So 2013; Mayes et al. 2014). In the international literature, one finds considerable numeracy research in countries such as Australia where numeracy is part of the adopted national curriculum standards and the United Kingdom that has a national numeracy educational initiative and testing program. The international body of work addresses numeracy development in the context of teacher education (e.g., Lucas et al. 2005; Pierce et al. 2013), teachers' practice, and use of numeracy tests results (e.g., Golding, 2001; McNamara and Corbin 2001; Carmichael et al. 2010;

Yang and Wu 2010; Aubrey et al. 2012; Pierce et al 2013). Notably absent are investigations into the development of numeracy beyond kindergarten and the early elementary years. In order to close the gap in our understanding of how numeracy progresses, and impacts children's academic, social and civic development, researchers need to have a psychometrically rigorous instrument with which to assess numeracy in the upper elementary years on into the high school years.

Background

For purposes of the present work, numeracy is defined as critical thinking applied to a quantitative context. Highly numerate individuals are skilled at applying their analysis, inference and interpretation skills in framing quantifiable data or making decisions using quantitative or probabilistic information. From this perspective, when we talk about improving numeracy what we are talking about is building critical thinking skills in the context of quantitative or probabilistic information (Steen 2000c). The critical thinking literature can be consulted, therefore, to theoretically ground and inform the assessment of numeracy.

Numeracy as Applied Critical Thinking

Critical thinking (CT) is widely recognized as an essential component of education and a powerful and vital resource in one's personal and civic life (Halpern 1996; Facione et al. 1997; Myers 2001; Winn 2004). In the past 25 years, educators and politicians have begun to seriously acknowledge the centrality of CT as an educational goal and outcome at the K-12 and postsecondary levels (Paul 1984; Ennis 1985; Lipman 1987; Facione 1990; Kuhn 1990; U.S. Department of Education 1990; Carter-Wells 1992; Halpern, 1996; Mayer 1997). There is great synergy between a focus on numeracy as critical thinking applied to the context of mathematics, probability and numerical data analysis and the overarching educational goal of building students' critical thinking in both the K-12 and postsecondary levels. Much like the numeracy literature, upper elementary and middle school students are under-represented in the domestic and international empirical literature on the development of critical thinking, with the exception of special populations such as gifted students (Diezmann and Watters 2001; Kettler 2014). Critical thinking is much more likely to be a searchable "key word" in studies of affiliate constructs such as creative thinking, philosophical inquiry, critical reflection, moral judgment, critical literacies. Frequently, investigations into critical thinking as the target educational outcome focus on teacher education programs or the attitude and behaviors of pre-service and in-service teachers (e.g., Eigenberger et al. 2001; Moseley and Ramsey, 2008; Sezar, 2008; Seda Saracaloglu et al. 2011).

Numeracy, or critical thinking in a quantitative context, emphasizes the use of analysis, inference, interpretation, explanation, evaluation, as well as reflection on one's own reasoning process (metacognition and self-regulation). Operationally, it refers to the ability to solve numerical and spatial-reasoning problems, draw inferences from quantifiable information in a variety of contexts, and reason probabilistically. Numeracy skills are used when applying one's knowledge of numbers, arithmetic, statistics, measures, and mathematical techniques to situations that require the interpretation and evaluation of quantitative information. Numeracy is more than being able to compute a solution to a mathematical equation. As Wiest et al. (2007) suggest, numeracy, or quantitative literacy, is not to be considered synonymous or subsumed by the discipline of mathematics. Instead, numeracy involves a recognition and understanding of how quantitative information is gathered, manipulated by counting and measuring, represented visually in graphs, charts, tables and diagrams and ultimately, dependent on context. Nor is numeracy redundant with statistics though their conceptual connection is not necessarily accidental. As explained by Steen (2000c), a joint campaign in the early 1990s between the American Statistical Association (ASA) and the National Council of Teachers of Mathematics (NCTM) to integrate elementary statistics and data analysis into the school curricula served as the origination of 'quantitative literacy' – a conscious, and in the end, non-limiting avoidance of the term statistics. Steen likens quantitative literacy, or its alternative expression numeracy, to a habit of mind and an approach to problem solving that employs the science of data (statistics) and the science of patterns (mathematics). Numeracy encompasses mathematical problem solving and decision making that goes beyond practical or daily-life pure-math applications to include civic, professional, leisure and cultural problems and decisions of varying novelty and complexity (Steen, 1990, 1999; Gittens and Facione 2014).

The highly influential resource publication, *Principles and Standards for School Mathematics* from the NCTM (2000) advocated a vision for school mathematics that included positioning students to be effective, skilled and reflective thinkers who can reason with evidence to confirm or disprove hypotheses and be resourceful problem solvers when working independently or within groups. This mathematics standards resource guide served as the foundation for the pre-K-8 curriculum focal points document (NCTM 2006) and the focus on reasoning and sense making in high school mathematics series (NCTM 2009) that emphatically affirmed the centrality of reasoning and problem-solving skills in pre-K-12 mathematics education. Not surprisingly, the national standards reform movement in mathematics, as reflected by the Common Core State Standards Initiative (CCSI) draws heavily on the NCTM vision and recommendations and prioritizes numeracy, particularly number sense and

problem solving, abstract and quantitative reasoning, argument construction and critique, structural analysis and strategic application of tools to solve math problems, and modeling with mathematics, as vital practice-based learning outcomes at all grade levels (Burns, 2012; CCSI 2014a). As of 2014, forty three states, the District of Columbia and four U.S. territories have adopted the Common Core Standards (CCSI 2014b). The widespread ratification and implementation of these academic standards has necessitated a commensurate instrument-development process for CCSI-aligned, large-scale, standardized assessment instruments. Most states have joined one of two state-led consortia, Partnership for Assessment of Readiness for College and Careers (PARCC) or the Smarter Balanced Assessment Consortia, who have been working to create comprehensive, CCSI standards-based instruments. The published timeline for full implementation of the PARCC and Smarter Balanced K-12 assessments for Mathematics and English Language Arts / Literacy is the 2014-15 academic year (PARCC 2014; Smarter Balanced 2014). Mathematical reasoning and problem solving figure prominently in the mathematics subtests of the PARCC and Smarter Balanced assessments. Given this imminent K-12 standardized testing and college-readiness assessment reform, educators and educational researchers alike would benefit considerably from the availability of focused numeracy scales that are conceptually aligned with the reasoning and problem-solving skills endorsed by NCTM and reflected in the language of the Common Core Standards for mathematics.

Purpose of the Current Study

There is clear direction coming from the National Council of Teachers of Mathematics (NCTM 2000) that K-12 mathematics education should focus on the development of reasoning and sense making. Similarly, there is a preponderance of critical thinking themes in the national and state-adopted Common Core Standards in Mathematics. These simultaneous and highly compatible thrusts in purpose suggest a national consensus that mastery in mathematics, namely the development of numeracy, depends on nurturing students' quantitative and data-analysis competencies in a learning environment that promotes students' critical thinking. The primary goal of this paper is to introduce a scale for assessing numeracy as an applied form of critical thinking among children and adolescents. The new numeracy scale is copyrighted by Insight Assessment, LLC (IA) and will be referred to as the IA numeracy scale. Proposed purposes of this instrument would be to enable educators and researchers to guide mathematics curriculum design and evaluation, shape pedagogical practices in order to differentiate instruction, to monitor development of numeracy skills at multiple points in time during the school year, to inform decisions about math placements, and to

maximize students' mathematics achievement. The development of this numeracy assessment occurred concurrently with the Common Core Standards Initiative, and the author closely monitored that national conversation for conceptual alignment. This instrument is not intended to be a measure of the breadth of content found in the Common Core Standards for Mathematics. This scale is intended to measure the underlying cognitive skill of quantitative reasoning that is integrated throughout these standards. Because numeracy has been likened to a habit of mind (Steen 2000c) or disposition, as well as a reasoning skill, a secondary intention was to share preliminary findings derived from frontier investigations of children's and adolescents' numeracy and critical thinking dispositional indicators.

Method

Participants

The participants in this cross-sectional study were 197 students (53% male, 47% female) from a private, Catholic, K-8 school in Santa Clara County, California. The school is one of the thirty K-8 schools in the Diocese of San Jose. All 234 students enrolled in grades 3 through 8 were invited to participate in the study in accordance with an approved protocol for the protection of human subjects. There was an 84% participation rate in this study. The racial / ethnic demographics of the school were: 46% White, 20% Hispanic, 13% Asian, 2% Black, and 19% reported being from two or more races.

Development of the Numeracy Scale

Numeracy has been operationally defined as the ability to critically reason about quantitative information. A fundamental priority was to isolate respondents' thinking by assuring that the items written for the IA numeracy scale developed in this project did not rely on mathematical computations. The goal was to develop items that would elicit the aforementioned skills of analysis, inference, interpretation, explanation, and evaluation, while minimizing or eliminating errors due to misapplication of formulas or other types of miscalculation. The target audience was upper elementary and middle school students because this was a population for which no numeracy instrument was available. The resulting new numeracy scale has two levels (or forms) in acknowledgement of the developmental span characteristic of children and adolescents within the targeted grade-level range. The IA numeracy scale is intended to be a stand-alone measure; nevertheless, it was anticipated that educators and educational researchers may wish to use it in conjunction with other instruments within a

single testing session. Therefore it was expected that both levels of the numeracy scale would be modest in length to minimize the risk of testing fatigue.

The IA numeracy scale was developed in three stages. The first stage was face-valid item creation. Research on children's arithmetic problem solving and algebraic reasoning, especially the work of Nathan and Koedinger (2000), provided a foundation for item construction in terms of quantitative reasoning. The draft Common Core Standards for Mathematics was referenced to identify developmentally appropriate mathematical content. Item content specifications for the IA numeracy scale are described below. A second stage focused on the refinement of item language in terms of understandability. The final stage was the pilot administration of items for scale refinement. The processes used to develop the instrument and to evaluate its reliability and validity are consistent with the *Standards for Educational and Psychological Testing* (AERA, APA and NCME 2014).

Stage One: Item Development. Literature on young children's arithmetic problem solving and algebraic reasoning (e.g., Riley et al. 1983; Kintsch and Greeno 1985; Carpenter et al. 1988; Carpenter et al. 1994; Carpenter and Levi 2000; Nathan and Koedinger 2000a, 2000b) was reviewed to provide the framework for item construction. Challenge and cognitive demand in relationship to numeracy was considered in terms of two important, structural factors that have been shown to influence item difficulty: (1) style of presentation used for the problem, and (2) the positioning of the unknown quantity in the problem (Nathan and Koedinger 2000a). The two familiar presentation styles are symbolic or computational problems and story problems. For example, a symbolic or computational problem would be solving for x in the following equation: $(3 \cdot 12) + 4 = x$. A comparable example of a story problem would be: *After the game Sophia gives one orange slice to each of her four soccer coaches and gives three slices to each of her 12 teammates. How many orange slices did Sophia bring to the game?* Nathan and Koedinger also describe a hybrid format, word-equations, that involve the description of relationships between quantities without the contextualizing scenario that characterizes the story-problem presentation style. An example of this would be: *Starting with 40, if I subtract 4 and then divide by 12, what number would I get?* Data on the impact of presentation style have not revealed a clear pattern of item response difficulties for younger student samples. Nathan and Koedinger's research (2000a; 2000b) suggested that among high school students, symbolic equation problems were more difficult than word equation and story problems. It is worth noting that the majority of secondary educators in Nathan and Koedinger's study predicted the exact opposite would be the case. The superior difficulty of symbolic equation problems has been replicated across diverse samples of 6th through 9th graders and adults (Nathan and

Koedinger 2000b). Verbal algebra problems were found to be just as challenging as symbolic arithmetic problems.

Positioning of the unknown quantity refers to whether the problem is asking for an unknown result (henceforth referred to as *result unknown*) that can be derived by performance of the mathematical operations identified in the problem or whether the respondent must derive a value or quantity necessary for specifying a mathematical relationship. The latter format is referred to as a *start-unknown* problem. If one considers the three examples of presentation offered above, it will be noted that these all reflect *result-unknown* positioning because the quantity being sought as a result is derivable from the operations included in the problem. These same examples can be modified to represent *start-unknown* problems with the following manipulations: the symbolic presentation would become $x \cdot 12 + 4 = 40$, the story problem would be *Sophia's dad prepared orange slices for each of her 12 teammates and one slice for each of her four soccer coaches. Sophia took a total of 40 orange slices to the game. How many slices did Sophia give each teammate?* And the word equation could be restated as *Starting with some number, if I multiply it by 12 and then add 4 I get 40. What number did I start with?* According to Nathan and Koedinger, research evidence consistently shows start-unknown problems to have a higher difficulty level than result-unknown problems (e.g., Riley et al. 1983; Koedinger and Tabachneck 1994; Koedinger and Nathan 2004; see Nathan and Koedinger 2000a for further discussion of this issue).

The sample question prompts offered above were adapted from the examples provided by Nathan and Koedinger (2000). They do not directly represent item formats used on the numeracy scale developed in this project. The IA numeracy scale items are not reproduced here because the scale is proprietary. Nevertheless, its items can be characterized generally and distinguished from the provided examples. One fundamental difference between the examples seen above and actual scale items is that, unlike the provided examples, items on the IA numeracy scale do not ask respondents to calculate a numerical answer. Therefore no symbolic equations were included. Rather, scale items require respondents to apply their critical thinking to either determine the optimal solution strategy or draw a reasonable conclusion from information provided. To clarify this distinction, the story-problem example could be modified to elicit evidence of a respondent's thinking skills. Instead of soliciting a calculation of the number of orange slices per teammate, the prompt could be rewritten to solicit an *explanation of the strategy necessary* for manipulating the provided information to ultimately derive the correct answer. The set up for the story-problem, result-unknown item prompt could include the statement *how can we figure out how many orange slices were given to each teammate?* The keyed (correct) response

for this item would be *subtract 4 from 40, then divide by 12*, and distractor responses might include choices such as *add 4 to 12 then subtract from 40*.

Quantitative reasoning item structure was derived from the literature described above. But reasoning takes place in context so the referential mathematical content needed to be determined. As was noted already, the development of the IA numeracy scale was underway at the same time as the national Common Core Standards were emerging. This new numeracy scale was intended to target upper elementary and middle school populations. As such, the Common Core Standards for Mathematics for grades 3 through 8 were monitored to ensure that the IA numeracy scale would reflect developmentally appropriate themes and procedural domains. The resulting Common Core Standards for Mathematics for grades 3 through 5 assert that students should be able to represent and interpret data; represent and solve problems using multiplication and division; engage in tasks requiring measurement, estimation, and pattern analysis; and write or interpret numerical expressions. The middle school years represent the transition from arithmetic to algebraic reasoning. In grades 6 through 8, students should be able to represent and solve arithmetic and algebraic expressions including one-variable equations; analyze quantitative and proportional relationships; engage in real-world problems requiring measurement, estimation, and functions; and reason inferentially about statistical variability, distributions and population data (CCSI 2014a). Success in these performance expectations requires numeracy.

The content and reasoning format considerations described above were strategically employed to guide the creation of a diverse item set. The items created for the two levels of the numeracy scale were written by the author and her research associates. The author and her colleagues have over three decades of recognized expertise and experience with critical thinking item writing and scale development for children, adolescents and adults. Approximately 30 items (15 items for each level) targeting upper elementary and middle school student populations were created, refined and ultimately reduced by this research team. The item specifications for reasoning format are presented in Table 1 and for mathematical content in Tables 2a and 2b.

Table 1.
Quantitative Reasoning Structural Format Classifications

	START UNKNOWN		RESULT UNKNOWN		TOTAL ITEMS ON SCALE
	Word Equations	Story Problems	Word Equations	Story Problems	
Upper Elementary Level	1	4	1	2	8
Middle School Level	1	6	2	2	11

Note. The classification system of item presentation style and positioning of unknown quantity is based on the research of Nathan and Koedinger (2000a).

Items on the IA numeracy scale were written in the presentation style and quantity-unknown location formats as described by Nathan and Koedinger (2000a), but purposefully included prompts to elicit thinking skills rather than computed results. Table 1 summarizes the number of items from the two structural factors: 1) start-unknown versus result-unknown position; and 2) word equations versus story problems for the two performance levels of the scale. On both levels there are more story problems than word problems. The purpose was to favor items that elicited numeracy as applied to real-world contexts. The IA numeracy scale contains more start-unknown items than result-unknown items. The focus on reasoning rather computation, and real-world scenarios rather than more-straightforward word equations, is consistent with authentic critical thinking conditions where a person is challenged to precisely understand the problem that she or he is trying to solve (Facione and Gittens 2013).

It was planned at the outset that the IA numeracy scale would have two levels (or forms) in acknowledgement of the wide developmental range characteristic of children and adolescents within the target grade levels. The two performance levels reflect differing degrees of cognitive complexity and challenge, but they target similar mathematical content domains. Because the quantitative reasoning and content-domain expectations broaden by grade level (CCSI 2014a), slightly more items were written for the middle school level of the scale.

Table 2a. Mathematical Content and Critical Thinking Skill Area Classifications: Upper Elementary Level

	ITEM RATIO	ANALYSIS	INTERPRETATION	EVALUATION	DEDUCTION	INDUCTION
Proportions / Ratio	4/8	100%	75%	50%	25%	75%
Measurement / Estimation	4/8	100%	50%	50%	0%	100%
Multiplication / Division	5/8	100%	40%	40%	40%	60%
Graphical Displays	2/8	100%	100%	100%	0%	100%
Arithmetic / Algebraic Expressions in Words	4/8	100%	75%	50%	75%	25%

Note. Item ratio represents the number of items from the total scale incorporating a specified mathematical content domain (rows). Cell percentages reflect the proportion of items requiring each specific critical thinking skill (columns) relative to the specified mathematical content domain.

The eight-item, upper elementary numeracy scale was created to engage students' reasoning in relation to the mathematical content and quantitative reasoning expectations endorsed by the Common Core Standards. The content domains are reflected in the left column of Tables 2a and 2b. The item-ratio column represents the number of numeracy scale items that encompass a specified mathematical content domain. For example, there are four items on the upper elementary level that require students to reason about proportions and ratios. It is important to note that items frequently engage students in more than one content domain. Looking again at the upper elementary level, there are five items where the respondents must reason about problems that employ multiplication or division. In some items, multiplication or division is contextualized within estimation or measurement scenarios. Other items present contexts where reasoning about multiplication or division is applied to proportions.

Table 2b. Mathematical Content and Critical Thinking Skill Area Classifications: Middle School Level

	ITEM RATIO	ANALYSIS	INTERPRETATION	EVALUATION	DEDUCTION	INDUCTION
Proportions / Ratio	6/11	100%	83%	100%	50%	50%
Measurement / Estimation	4/11	100%	50%	50%	0%	100%
Multiplication / Division	7/11	100%	57%	43%	57%	43%
Graphical Displays	2/11	100%	100%	100%	0%	100%
Arithmetic / Algebraic Expressions in Words	6/11	100%	83%	50%	83%	17%
One-variable Equations	1/11	100%	100%	100%	100%	0%

Note. Item ratio represents the number of items from the total scale incorporating a specified mathematical content domain (rows). Cell percentages reflect the proportion of items requiring each specific critical thinking skill (columns) relative to the specified mathematical content domain.

The remaining columns on these tables identify the reasoning skills that are engaged. Definitions of the analysis, interpretation and evaluation skill areas are based on the work of Facione and colleagues since the publication of the APA expert consensus definition of critical thinking (Facione 1990; Facione et al. 2014). *Analysis* refers to the process of identifying assumptions, reasons and claims, gathering detailed information from charts, graphs, diagrams as well as narrative text, and to closely examine the collected ideas and information.

Interpretation refers to the process of determining the exact meaning of a word, phrase, text, idea, claim, sign, graph, diagram, or chart in a given context for a specific purpose. The cognitive skills involved in interpretation include categorizing information and decoding the significance and clarifying the meaning of information. *Evaluation* involves assessing the credibility of claims and the strength or weakness of arguments. One can also be engaging in evaluation when forming judgments about the quality of inferences, analyses, interpretations, explanations, opinions, claims, proposals and conclusions.

The final two columns classify items in the more-traditional categories of deductive and inductive reasoning (Facione et al. 2014). *Deduction*, or axiomatic reasoning, is the process of drawing valid inferences. A valid deductive argument requires that the conclusion must be true if the premises are all true. Algebra, geometry, many computer programs and even Sudoku puzzles are examples of deductive reasoning because they involve the precise following of rules, definitions, laws or principles. *Induction*, or probabilistic reasoning, is drawing a warranted inference about what is most likely true or most likely not true, given the information and context under consideration. The premises of an inductive argument are taken to be true on the basis of verifiable evidence. Information that goes into inductive reasoning comes from data, key examples, precedents, past experiences, or relevant cases. As long as there is a chance that the conclusion might be inaccurate, the thinking process is inductive. Probabilistic reasoning is inductive in character, open to self-corrective revision.

Cell percentages in Tables 2a and 2b reflect the proportion of items requiring a given critical thinking skill relative to the specified mathematical content domain. Each item requires more than one critical thinking skill to determine the correct answer. Analysis must lead to interpretation or evaluation, or both, to formulate a response to a test item. One should note, however, that items require deductive reasoning or inductive reasoning but never both. Every item on our numeracy scale requires students to analyze the information provided in the question prompts. Variability in percentages in a given row would be interpreted in the following way. In Table 2b, for example, there are four (out of 11) items on the middle school level that engage students in reasoning in the context of measurement or estimation. Of those four items, all require analysis, half require interpretation, and half require evaluation. All four items require inductive reasoning.

Each of the eight multiple-choice items on the upper elementary level was accompanied by four response options (labeled A-D) that included the keyed (correct) answer and three closely related but incorrect distractors. Each of the 11 multiple-choice items on the middle school level was accompanied by five response options (labeled A-E) that included the keyed answer and four closely related but incorrect distractors.

Stage Two: Item Refinement and Administration. A think-aloud activity with a small number of children from grade levels 3-8 was performed to confirm that students engaged the appropriate cognitive skills when attempting to answer the test items. These interviews were helpful in reviewing the wording and understandability of the instructions and item prompts. Students who had received parental consent to participate in the study were invited to volunteer for this item-refinement session. Interviews with two students representing each target grade-level range were conducted by the researcher. The most-common feedback from these sessions was that the questions were interesting and fun to figure out. The interviews did produce recommended edits from the younger students to simplify the instructions for recording answers on the answer sheet on the upper elementary numeracy scale. These were the only refinements made at this stage.

Stage Three: Pilot Testing of the Two Level Numeracy Scale. To conduct pilot testing, the two versions of the numeracy scale were included in a larger, established, critical thinking battery (the *California Critical Thinking Skills Test-M Series*, Facione et al. 2014, described below). The critical thinking instrument with the embedded numeracy scale was administered in the fall of 2010 to consenting students from 3rd through 8th grade classrooms at the participating school. An average of 30 students per classroom participated in this stage. The researcher and two trained undergraduate assistants circulated among grade-level groups to answer student questions during the testing session. Whenever students sought clarity of instructions or the meaning of words or phrases, they were instructed to circle words or phrases they found challenging or confusing. On average, less than three questions were asked in a given classroom, and these questions were mostly from students who were seeking affirmation that they had selected the correct answer rather than wording or instructional clarifications. Evaluation of students' questions during the testing session as well as a review of written notations on the test booklets resulted in only minor edits to improve readability. After this pilot administration the items included on the instrument (two levels) were finalized.

Measurement Instruments

Numeracy was evaluated in the present study as an embedded scale within a broader, established assessment of general critical thinking skills. The critical thinking skills instrument is available in two levels, one designed for upper elementary students and the other for middle school students. A critical thinking dispositions questionnaire was also administered as a means for investigating the attitudinal correlates of numeracy. School site archival data related to cognitive abilities and math academic achievement were compiled for purposes of construct validation.

The California Critical Thinking Skills Test - M Series: Versions M1B and M25. Items forming the two levels of our numeracy scale were embedded within the *California Critical Thinking Skills Test – Middle School level* (CCTST-M25; Facione et al. 2014) and the *California Critical Thinking Skills Test – Upper Elementary level* (CCTST-M1B; Facione et al. 2014). The upper elementary school version of the CCTST was first introduced in 2005, and it was followed in 2008 by the introduction of the middle school version. The CCTST family of instruments, which includes levels for older students and adult samples, is based on the Delphi study sponsored by the American Philosophical Association (APA) that produced an expert consensus definition of critical thinking as *the process of purposeful, self-regulatory judgment* (Facione 1990). Test items on the CCTST instruments employed in this study use everyday scenarios that are common and familiar to children and adolescents. Informational content needed to respond correctly is provided within each stem. Items on the CCTST instruments require test takers to analyze the scenario information, draw appropriate inferences, and evaluate the arguments presented. Performance on the CCTST is represented by an overall scale score as well as subscale scores. The CCTST instruments, like other nationally normed, standardized tests, are not static. The master pool of items is frequently revised and replenished to retain the currency of real-world problem-solving scenarios. The CCTST instruments are periodically re-versioned and re-normed to maintain instrument security while not compromising statistical equivalence across iterations. Since its development, the CCTST subscales have included the traditional logic scales of inductive and deductive reasoning, as well as scores on analysis and interpretation, inference, and evaluation and explanation. The IA numeracy scale represents a formal expansion of the *CCTST M Series*. Critical thinking, as defined by the APA Delphi study, is an interactive, iterative process comprising these underlying cognitive skills. The items on the subscales often engage multiple CT skills. Scales on the *CCTST M Series* instruments are calculated using a weighted combination of shared and unique items. Due to the statistical dependence among the CCTST subscales, inter-scale correlations matrices are not reported. For these reasons, the CCTST was not used to evaluate the validity of the numeracy scale (AERA, APA and NCME 2014).

IA Numeracy Scale. The IA numeracy scale was administered as an embedded measure on the *CCTST M Series* instruments. The inclusion of a numeracy scale, as a sixth subscale, is a natural extension of the assessment of critical thinking in the applied domain of quantitative reasoning and data analysis. Though administered as an embedded measure here, the IA numeracy scale is available for use as a stand-alone scale. Numeracy skills are used when applying knowledge of numbers, arithmetic, measures, probabilities, mathematical and data analytic techniques to situations that require the interpretation or evaluation of information. Items on the numeracy scale require respondents to demonstrate

understanding of how quantitative information can be manipulated by counting and measuring, used for prediction (probability), represented visually such as in graphs, charts, tables and diagrams, and used to solve real-world problems.

A readability analysis was performed on both levels of the newly developed, embedded numeracy scale using the Flesch-Kincaid Grade Level formula available in MSWord. The 8-item multiple-choice numeracy scale for upper elementary students had a Flesch-Kincaid Grade Level index of 4.2, and the 11-item numeracy scale for middle school students had a Flesch-Kincaid Grade Level index of 6.8. Though this formula is considered a coarse estimate of reading level (Begeny and Greene 2014) these results suggest that the readability levels are suitably matched to the targeted grade levels. The reliability estimates for the levels of the numeracy scale were evaluated. A Kuder-Richardson Formula 20 (KR-20; Kuder and Richardson 1937) was used to determine the level of agreement between items. The KR-20 coefficient for the upper elementary level of the numeracy scale was .40 and for the middle school numeracy scale was .57. These observed values are below the recommended level of .70 which is considered a more acceptable standard for dichotomous scales (see Thompson 2003). It is hypothesized that reliability estimates in the current samples were adversely affected by the number of items per scale, sample size, range restriction (Fife et al. 2012), and violation of the tau-equivalent measurement model that assumes that all items measure the same latent construct (Komaroff 1997; Graham 2006). The Spearman-Brown prophecy formula (Wells and Wollack 2003) was used to determine the effect on reliability that could be anticipated by adding like items to the numeracy scale. To achieve the standard of a $KR-20 \geq .70$, the upper elementary level would need 28 items with the current sample size. The new reliability coefficient for the middle school level would be .73, if the number of items was doubled to 22. The numeracy scale is intended as a stand-alone instrument that could be administered in conjunction with other measurement instruments during a single administration session. The benefit of reducing measurement error by increasing the number of items on the scale must be weighed against that goal. Evidence suggests, however, that expanding the scale may be a warranted action. Inadequacy of sample size was also proposed as an explanation for the observed reliability coefficients. As partial confirmation that sample size in the current study was insufficient to confirm the stability of measurement, the author was granted access by the test publisher to a diverse, anonymous, aggregated sample (N=305) of middle school-level numeracy scale data. In this larger, independent sample a KR-20 of .69 was observed. Sample-size sensitivity aside, future investigations of scale stability are warranted.

California Measure of Mental Motivation (CM3). The California Measure of Mental Motivation (CM3; Gittens and Facione 2014) is an attitudinal survey designed to gauge respondents' disposition toward critical thinking. The phrase

critical thinking disposition refers to a person's internal motivation to think critically when faced with problems to solve, ideas to evaluate, or decisions to make (Facione et al. 2000; Giancarlo et al. 2004). These attitudes, inclinations, or habits of mind are dimensions of one's personality and motivational style which relate to how likely a person is to approach decision-making contexts or problem-solving situations by using their reasoning skills. The CM3 is available in four developmental levels: Level IA for grades Kindergarten through 2nd (primary), Level IB for grades 3-5 (upper elementary), Level II+ for middle school and secondary students, and Level III for post-secondary and adults (Gittens and Facione 2014). Levels IB and II+ were administered in the present study. The CM3 IB is a 25-item survey that employs a dichotomous response format, agree or disagree. The CM3 II+ is a 72-item survey that employs a four-point Likert-type scale ranging through agree strongly, agree, disagree or disagree strongly. The CM3 produces scores on four dispositional dimensions: Learning Orientation, Creative Problem Solving, Mental Focus, and Cognitive Integrity (Giancarlo et al. 2004; Giancarlo-Gittens 2010). Descriptions of these scales and reliability estimates generated from the current samples are presented in Table 3. Alpha reliability statistics computed for the four scales of the CM IB ranged from a low of .41 to a high of .79. Reliability indices for the considerably longer CM3 II+ instrument ranged from a low of .71 to a high of .83.

Table 3.
CM3 Scale Descriptions with Reliability Estimates

SCALE	DESCRIPTION	VERSION	RELIABILITY (α)
Learning Orientation	The inclination or desire to increase one's knowledge and skill base.	IB	.41
		II+	.75
Creative Problem Solving	The tendency to approach problem solving with innovative or original ideas and solutions.	IB	.79
		II+	.81
Mental Focus	Self-described diligence, focus, systematicity, task-orientation, organization, and clear-headedness.	IB	.51
		II+	.71
Cognitive Integrity	The motivation to use one's thinking skills in a fair-minded fashion.	IB	.68
		II+	.83

Note. N= 93 for the CM3 Level IB (grades 3-5) and N = 87 for the CM3 Level II+ (grades 6-8).

Iowa Test of Basic Skills (ITBS). *The Iowa Test of Basic Skills* (Hoover et al. 2001) is a well-established, standardized, paper-and-pencil, academic achievement test battery developed by the Iowa Testing Programs at the University of Iowa. The ITBS is geared for grades K-8; it is available to schools and districts in ten articulated levels, numbered 5 through 14 (Riverside Publishing 2010a). The first four levels are appropriate for grades K-2, and Levels 9-14 are designed for administration in 3rd through 8th grade. The ITBS Levels 9-14 include scales representing an array of subject areas: mathematics, vocabulary,

reading comprehension, language, social studies, and science. The Math test from the ITBS Levels 9-14 was used in the present study. According to the test publisher, the ITBS mathematics test measures students' ability "to do quantitative reasoning and to think mathematically in a wide variety of contexts" (Riverside Publishing 2010a). The three subscales of the ITBS Math test, Math Concepts and Estimation, Math Problem Solving and Data Interpretation, and Math Computation, were recorded for each participating student. Kuder-Richardson reliability coefficients for the ITBS mathematics subtests have been reported in the range of .70-.90 (Iowa Testing Programs 2014). Computational skills are required on all three subscales, but they are most prevalent on the math computation subscale. The ITBS was included in the present study because it represented an independent assessment of quantitative reasoning and includes items that focus on the understanding of operations and the interpretation of data displays on tables and graphs. The ITBS is a familiar alternative used in private schools that do not administer the achievement tests that are required of all public schools in the state of California. The ITBS battery is administered annually in the fall term at the school where this investigation occurred.

Cognitive Abilities Test (CogAT). *The Cognitive Abilities Test* (Lohman and Hagen 2001) measures reasoning ability in three domains: verbal, quantitative, and nonverbal. Each of the three scales is measured through a set of subtests that are typically administered over three different sessions. According to the test publisher, the CogAT when paired with the ITBS can be used for placement decisions or to enhance the educational experience of all students (Riverside Publishing 2010b). The CogAT was included in the present study because it specifically targets students' reasoning skills as well as includes a quantitative scale. These characteristics make the CogAT well suited as a measure of construct validity for the IA numeracy scale. The CogAT is administered bi-annually by the school in the spring term to all 3rd, 5th, and 7th graders. Thus, CogAT data were available from approximately half of the total sample.

Procedure

Data were collected during the spring of 2010 and the 2010-11 school year. The timeline for scale development and data collection is presented in Table 4. The CCTST skills instrument with the embedded numeracy scale was administered in the fall to pilot the new scale and then again in the spring for the validation study. The CM3 dispositional measure was administered during both the fall and spring terms though fall term CM3 data were not used in this paper. Teachers for grades 3 through 8 (six teachers, one per grade level) assisted with the administration of the critical thinking instruments in fall and spring. The researchers and trained

research assistants were on site during these administration sessions to provide assistance as needed. In independent testing sessions that year, the ITBS and the CogAT were administered by school personnel as part of the achievement testing program conducted throughout the Diocese. Student achievement results were collected from students' cumulative files at the end of the school year. For all students, this included scores on the mathematics scale and associated subscales from the ITBS academic achievement test. For students in grades 3, 5, and 7 this also included composite and scale scores from the CogAT. Because data were collected during two time frames in the academic year, the sample sizes in any given statistical evaluation reflect the variability in number of students present on the day the study instruments or standardized tests were administered.

Table 4.

Tool Development and Data Collection Timeline

FALL 2009	SPRING 2010	FALL 2010		SPRING 2011	
Stage 1: Item Development	Stage 2: Item Refinement Numeracy Scale only	Stage 3: Pilot CCTST-M Series with Numeracy Scale; CM3 critical thinking dispositional measure	ITBS Testing (School-wide)	CCTST-M Series with Numeracy Scale; CM3 critical thinking dispositional measure	CogAT Bi-annual Testing (3 rd , 5 th and 7 th graders only)

Note. Data collected in fall 2010 and spring 2011 are presented in separate columns to represent separate occasions; ITBS and CogAT testing was performed by school officials independent of this validation study.

Results

The primary focus of the present study was to gather support for the construct validity of a numeracy scale designed for students at the upper elementary and middle school level. The evaluation of construct validity, however, was necessarily hampered by the dearth of numeracy measurement instruments designed for the target population. The numeracy scale discussed in this report is offered as a means for closing the gap in the assessment literature. The construct validity of the IA numeracy scale was examined with respect to two published measures of mathematical achievement. A significant positive association with the ITBS Total Math scale and all three subscales was anticipated. Because the IA numeracy scale targets quantitative reasoning rather than computational skills, it was expected that numeracy would have the weakest, yet still substantial, association with the Math Computation subscale. It was predicted that numeracy would be more strongly associated with the Math Concepts and Estimation subscale. The strongest association was anticipated between numeracy and the Problem Solving and Data Interpretation subscale. A non-zero correlation was expected with all three subscales of the CogAT because it is a reasoning ability test. It was hypothesized that the IA numeracy scale would have the greatest

association with the Quantitative scale due to the conceptualization of numeracy as critical thinking applied to a quantitative context. Weaker but still significant associations were expected with the Verbal and Non-verbal reasoning scales, with greater association anticipated with verbal reasoning due to the word equation and story problem delivery format of the numeracy items.

A secondary aim was to explore numeracy as a habit of mind (Steen 2000c). The Learning Orientation, Creative Problem Solving, Mental Focus and Cognitive Integrity scales of the CM3 collected during the spring administration session were used in the exploratory review of dispositional associations. Results are presented separately for the two developmental levels of the numeracy scale; whenever possible, results were also disaggregated by grade level. Because data from the fall administration of the CCTST with the embedded numeracy scale were used for purposes of Stage Three pilot testing of the scale, the results for construct validity are based on data collected during the spring administration session.

Construct Validity of the Numeracy Scale

Numeracy scores were strongly and positively correlated with the ITBS Math Total score and three subscales in the predicted direction and relative strength. These relationships were found for the two developmental levels of the numeracy scale as well as when the data were disaggregated by grade levels. Results are presented in Table 5. The numeracy scale for upper elementary was correlated .680 and the scale for middle school was .574 with ITBS Math Total scores. This suggests that 46% of the variance in ITBS overall math achievement for the upper elementary sample and 33% of the variance in the ITBS overall math achievement for the middle school sample was explainable in terms of numeracy as measured by this new instrument. The individual grade level correlations ranged from .421 to .779, with an average correlation of .643 with ITBS Math Total. Correlations in this sample were stronger among the younger students despite lower reliability and fewer items on that version of the numeracy scale. As predicted, students' numeracy scores and their performance on the Math Concepts and Estimation and the Math Problem Solving and Data Interpretation subscales were also strongly, positively correlated across grade levels in the .5 to .8 range. The lower correlations with the math computation subscale can be attributed to the fact that this ITBS scale consists particularly of problems that require students to solve arithmetic equations and rely less on reasoning than the other two subscales. This pattern of correlations is consistent with the proposed construct of numeracy as it informed the development of this scale. Also, these results suggest that the numeracy subscale of the critical thinking instruments has independent predictive validity for math achievement.

Table 5.
Numeracy Scale Correlations with ITBS Mathematics Achievement

	N	Math Total	Concepts/ Estimation	Problem Solving/ Data Interpretation	Computation
Numeracy: Upper Elementary	85	.680	.645	.705	.497
Grade 3	28	.711	.719	.601	.558
Grade 4	26	.605	.494	.641	.441
Grade 5	31	.779	.721	.814	.503
Numeracy: Middle School	72	.574	.568	.623	.395
Grade 6	16	.770	.696	.763	.618
Grade 7	27	.576	.604	.588	.380
Grade 8	29	.421	.389	.543	.091

Note. The Upper Elementary Numeracy Scale was embedded in the CCTST-MIB and the Middle School Numeracy Scale was embedded in the CCTST-M25.

As can be seen in Table 6, students' numeracy scores for both levels were strongly and positively associated with the CogAT composite score, and the three scales of the CogAT. Correlational results disaggregated by grade level are also presented in Table 5. These correlations indicate that 40 to 56% of the variance in the CogAT composite scores was explainable by scores on the numeracy scale. Numeracy was strongly and positively correlated with the Quantitative CogAT scale scores in the sample of 3rd, 5th and 7th grade students with a range of .661 to .726. Here, too, 43 to 53% of the variance in the CogAT Quantitative scale scores was explainable in terms of the numeracy measure.

Table 6.
Numeracy Scale Correlations with the Cognitive Abilities Test Subscales

	N	Quantitative	Verbal	Non-Verbal	Composite
Grade 3	29	.661	.606	.349	.636
Grade 5	30	.712	.663	.604	.748
Grade 7	29	.726	.569	.424	.677

Note. Time2 (spring) numeracy scale scores were used for this analysis. Students in grades 3-5 completed the upper elementary level of the numeracy scale embedded on the CCTST-MIB and students in grades 6-8 completed the middle school level of the numeracy scale embedded in the CCTST-M25.

Numeracy is measured using word and story problems that must be read, so a strong correlation between the numeracy scale and measures of verbal ability was predicted. Furthermore it was anticipated that the association between numeracy and CogAT verbal reasoning would be stronger than the association with nonverbal reasoning. Neither set of correlations was expected to be as strong as the association found with the quantitative CogAT scale. As predicted, the correlations between numeracy and the CogAT Verbal scale ranged between .57 and .66 (explaining 32 to 44% of the variance) and between .349 and .604 (explaining 12 to 36% of the variance) with the Non-verbal scale. These findings support the conclusion that the numeracy scale assesses a students' ability to reason about information – in this case, quantitative information – and it is not an assessment of students' ability to solve mathematical equations as a proxy indicator of quantitative reasoning ability. The strong positive correlations observed between the numeracy scale and the CogAT scales, especially the quantitative cognitive abilities subscale, are consistent with the hypothesis that a meaningful, predictive relationship exists between these constructs.

Taken together, these correlational results support the conclusion that numeracy skills as assessed by the IA numeracy scale are consistently, strongly, and positively related to students' math achievement and quantitative, verbal and non-verbal cognitive abilities. Furthermore, these relationships have been documented at multiple grade levels ranging from early elementary to middle school. It is important to note that the samples, especially when disaggregated by grade level, are small in this study; thus, these correlations may not be stable. Replication studies with larger samples will allow the stability of these relationships to be evaluated. With tentative confidence in the construct validity of the new numeracy instrument, the analyses turned to the more exploratory question of whether numeracy is related to critical thinking dispositions.

Dispositional Correlates of Numeracy

Numeracy scores were correlated with the four scales of the CM3 and the results are presented in Table 7. Prior to these analyses, scatter plots of the dispositional scales were viewed and outliers were removed. Among the younger students, numeracy was found to be positively associated with creative problem solving, mental focus, and cognitive integrity but not with learning orientation. These variables were not significantly associated in the older-student sample. An initial interpretation of these results is offered in the section below. As was stated above, these findings are preliminary and additional samples of greater size will be necessary to investigate these potential associations in sufficient depth and detail.

Table 7.
Numeracy Correlations with Critical Thinking Dispositional Scale Scores on the CM3.

	N	Learning Orientation	Creative Problem Solving	Cognitive Integrity	Mental Focus
Version 1B	84	.199	.353**	.332**	.348**
Grade 3	28	.154	.192	.346	.358
Grade 4	26	.170	.425*	.223	.554**
Grade 5	30	.325	.465**	.283	.250
Version II+	81	-.157	.117	.050	.092
Grade 6	17	-.243	.222	.359	.054
Grade 7	29	-.236	-.113	-.179	-.005
Grade 8	31	.013	.232	-.039	.215

Note. These numeracy and CM3 data were collected in the spring of 2011. * indicates statistical significance at the $p < .05$ level. ** indicates statistical significance at the $p < .01$ level.

Discussion

The goal of the present study was to describe a new measure of numeracy aligned with an emerging international conceptual consensus for use with upper elementary and middle school students. The intention of the researcher was to provide a valid mechanism for assessing quantitative reasoning among a student population that is currently underrepresented in the numeracy literature. The eight items of the numeracy scale for upper elementary and the eleven items of the numeracy scale for middle school cohorts provide an assessment that is conceptually aligned with the reasoning and problem-solving skills reflected in the Common Core Standards for mathematics and the body of empirical literature on the development of children's arithmetic problem solving and algebraic reasoning. Findings from this investigation support the tentative conclusion that the IA numeracy scale is gathering validity evidence as an assessment of quantitative reasoning for use with the targeted student populations. As hypothesized, a strong positive relationship was demonstrated between students' numeracy and the mathematics domain of the ITBS, a widely used standardized measure of academic achievement, as well as the CogAT, an established indicator of quantitative, verbal and nonverbal cognitive abilities. These strong positive correlations were observed at all grade levels included in the study. These results are provocative in that they affirm, through empirical evidence, the statistical and practical significance of nurturing students' numeracy in order to enhance mathematics achievement and cognitive abilities in both the quantitative and the verbal and nonverbal domains. To that end, educators and researchers may consider use of the IA numeracy scale at the classroom, school, or district level to

establish baseline performance and monitor growth in quantitative reasoning skill development at multiple points in the academic calendar.

It has been argued that numeracy or quantitative literacy is a powerful moderator for informed and engaged citizenship (Steen 2000b; Root 2009). For example, Root (2009) has asserted that quantitative literacy is an essential tool for appreciating societal constructs such as social justice and the equitable distribution of benefits and burdens among the individuals living in community. Wiest et al. (2007) offer a stronger perspective suggesting that quantitative literacy is an essential gatekeeper for effective functioning in today's society and therefore the development of quantitative literacy among the citizenry should be viewed as a matter of social justice in the Information Age.

A key element of this educational commitment is the nurturing of both thinking skills and thinking attitudes, or dispositions. It is not sufficient for educators to nurture students' numeracy skills if, when faced with a problem that demands quantitative reasoning, the student fails to exercise what they have learned. For this reason, an investigation was undertaken to explore the dispositional correlates of numeracy. The results suggest that there is an association between numeracy and creative problem solving, mental focus and cognitive integrity, at least among the younger students. These findings are consistent with cognitive science research that continues to clarify the relationship between attention and the creative process during problem solving (see Mendelsohn 1976; Simon and Reed 1976; Peterson et al. 1984; Kasof 1997; Runco and Sakamoto 1999; Vartanian et al. 2007; Vartanian 2009; Yang and Chang 2013).

Limitations

The current study provides tentative evidence of the validity of a numeracy scale for upper elementary and middle school students, thus filling a critical gap in the numeracy assessment literature. However this study is not without limitations. The sample of students was arguably small, especially when disaggregated at the grade level. Furthermore, it was from a private, co-educational, faith-based, K-8 school in California. While the gender distribution was balanced and the ethnic / racial diversity of the sample was acceptably heterogeneous, the potential compromise to generalizing to public school samples or to different geographic regions warrant mention. It was posited that these limitations of sample size and range restriction may be contributing factors to the low reliability estimates observed. Because validity coefficients are limited by the reliability of the testing instruments in a given study, the observed results should be granted appropriate skepticism pending replication. Continued research is called for to strengthen the internal consistency of the scale, as well as extend validity investigations to include larger and differentiated student samples.

The lack of association between numeracy and learning orientation in the present sample warrants comment. It is important to note that the reliability of the learning-orientation scale among the upper elementary students was disappointingly low compared to the other three scales on the CM3 IB instrument. A review of students' scores on the Learning Orientation scale revealed a notable ceiling effect, so it may not have been an appropriate study criterion. This could explain the lack of significant relationship for this variable. The situation becomes more curious, however, when one ponders the different profile of outcomes that is revealed when the results for the middle school students fail to mirror those of the upper elementary students. A review of the published empirical literature on students' learning orientation specifically, or academic motivational goals more broadly, in the mathematics domain reveals complex and varied associations with positive relationships more likely to be found with younger samples than older samples. For example, researchers have found a positive relationship between teachers' ratings of preschoolers' motivation and their counting and spatial skills (Edens and Potter 2013). Among upper elementary students, those who are challenge seekers and have a learning-goal orientation are more likely to have higher self-efficacy in math (Meyer et al. 1997). Studies of middle school students however have shown that goal orientation was related to performance for males only (Patrick et al. 1999) and yet others report shifting grade level and math-type associations among middle school students (Cleary and Chen 2009). More research into the connection between numeracy and other dispositional and motivational factors is needed to reveal the nuances of this relationship over the school years. It stands to reason that only through the combined effort to promote numeracy while nurturing the desire to be a confident and capable thinker will we produce future generations of leaders who will be capable of solving the significant global challenges like climate change, poverty, and HIV/AIDS that define our technological age. This line of reasoning is what prompted the exploratory question in the current study and is aligned with similar inquiries into the role of thinking dispositions in the development of quantitative skills in the sciences as part of science learning (Quinnell et al. 2013).

Concluding Remarks

Numeracy has been defined in this paper as critical thinking applied to a quantitative context. If we are to accept the argument that numeracy rivals reading literacy and language fluency in its importance as a vital life skill and habit of mind, it should go without question that the development of numeracy should share the same platform as critical thinking as a fundamental outcome of K-12 education. In short, the educational system should produce graduates who are willing and able to use their cognitive powers of analysis, interpretation,

inference, evaluation, explanation, and self-monitoring meta-cognition to make purposeful judgments about quantitative and probabilistic information as it pertains to what we believe or do. The current empirical literature has heretofore been without a valid measure of numeracy for students in the upper elementary and middle school years. With the introduction of the numeracy scale in the current study, researchers will now be able to connect the research trajectories of numeracy development in early childhood and the manifestations of numeracy among post-secondary and adult samples.

Acknowledgments

The author wishes to thank the students, teaching faculty, site administrative staff, and school community who participated in this study. The author acknowledges the undergraduate student workers at Santa Clara University who assisted with data collection.

References

- AERA, APA and NCME. See American Educational Research Association, American Psychological Association, National Council on Measurement in Education. American Educational Research Association, American Psychological Association, National Council on Measurement in Education, Joint Committee on Standards for Educational and Psychological Testing (US). 2014. *Standards for educational and psychological testing*. American Educational Research Association.
- Aubrey, C., K. Ghent, and E. Kanira. 2012. Enhancing thinking skills in early childhood. *International Journal of Early Years Education* 20(4): 332–348. <http://dx.doi.org/10.1080/09669760.2012.743102>
- Begeny, J. C. and D. J. Greene. 2014. Can readability formulas be used to successfully gauge difficulty of reading materials? *Psychology in the Schools* 51(2): 198-214. <http://dx.doi.org/10.1002/pits.21740>
- Burns, M. 2012. Go figure: Math and the Common Core. *Educational Leadership* 70(4): 42–46.
- Carmichael, C., Callingham, R., Hay, I. and J. Watson. 2010. Statistical literacy in the middle school: The relationship between interest, self-efficacy and prior mathematics achievement. *Australian Journal of Educational & Developmental Psychology* 10: 83–93.
- Carpenter, T. P. and L. Levi. 2000. *Developing conceptions of algebraic reasoning in the primary grades*. Report No. 00-2. Madison: Wisconsin Center for Educational Research (NCISLA).
- Carpenter, T. P., E. Fennema, P. L. Peterson, and D. A. Carey. 1988. Teachers' pedagogical content knowledge of students' problem solving in elementary arithmetic. *Journal for Research in Mathematics Education* 19(5): 385–401. <http://dx.doi.org/10.2307/749173>

- Carpenter, T. P., E. Fennema, and M. L. Franke. 1994. *Cognitively guided instruction: Children's thinking about whole numbers*. Madison: Wisconsin Center for Education Research.
- Carter-Wells, J. A. 1992. Defining, teaching, and assessing critical thinking in a multicultural context. *Association of American Colleges Annual Meeting*, Washington D.C.
- CCSI. See Common Core Standards Initiative.
- Chan, W. W., T. K., Au, and J. Tang. 2013. Developmental dyscalculia and low numeracy in Chinese children. *Research in Developmental Disabilities* 34(5): 1613–1622. <http://dx.doi.org/10.1016/j.ridd.2013.01.030>
- Cleary, T. J. and P. P. Chen. 2009. Self-regulation, motivation, and math achievement in middle school: Variations across grade level and math context. *Journal of School Psychology* 47(5): 291–314. <http://dx.doi.org/10.1016/j.jsp.2009.04.002>
- Common Core State Standards Initiative. 2014a. *Common Core Standards for Mathematics*. Retrieved from <http://www.corestandards.org/Math/>.
- . 2014b. *Standards in Your State*. Retrieved from <http://www.corestandards.org/standards-in-your-state/>.
- Crowe, A. R. 2010. “What’s math got to do with it?”: Numeracy and social studies education. *The Social Studies* 101: 105–110. <http://dx.doi.org/10.1080/00377990903493846>
- Diezmann, C. M. and J. J. Watters. 2001. The collaboration of mathematically gifted students on challenging tasks. *Journal for the Education of the Gifted* 25(1): 7–31.
- Edens, K. and E. Potter. 2013. An exploratory look at the relationships among math skills, motivational factors, and activity choice. *Early Childhood Education Journal* 41(3): 235–243. <http://dx.doi.org/10.1007/s10643-012-0540-y>
- Eigenberger, M. E., K. A. Sealander, J. A. Jacobs, and S. M. Shellady. 2001. Dispositions toward thinking critically: A comparison of pre-service teachers and other university students. *North American Journal of Psychology* 3(1): 109–119.
- Ennis, R. 1985. The logical basis for measuring CT skills. *Educational Leadership* (43)5.
- Facione, P. A. 1990. Critical thinking: A statement of expert consensus for purposes of educational assessment and instruction. *The Delphi Report: Research findings and recommendations prepared for the committee on pre-college philosophy*. American Philosophical Association. (ERIC Doc. No. ED 315 423).
- Facione, P. A. and C. A. Gittens. 2013. *THINK Critically*, 2nd Ed. Upper Saddle River, NJ: Pearson.
- Facione, P.F., N. C. Facione, and C.A. Giancarlo. 1997. The motivation to think in working and learning. In *Preparing competent college graduates: Setting new and higher expectations for student learning*, ed. E. A. Jones, 67–79. San Francisco: Jossey-Bass.
- . 2000. The disposition toward critical thinking: Its character, measurement and relationship to critical thinking skill. *Informal Logic* 20: 61–84.
- Facione, P. A., N. C. Facione, C. A. Gittens, and K. Winterhalter. 2014. *The California Critical Thinking Skills Test – M-Series Test Manual*. Millbrae, CA: Insight Assessment / California Academic Press. <http://www.insightassessment.com/>

- Fife, D. A., J. L. Mendoza, and R. Terry. 2012. The assessment of reliability under range restriction: A comparison of α , ω , and test-retest reliability for dichotomous data. *Educational and Psychological Measurement* 72: 862–888. <http://dx.doi.org/10.1177/0013164411430225>
- Foegen, A. 2008. Progress monitoring in middle school mathematics. *Remedial and Special Education* 29(4): 195–207. <http://dx.doi.org/10.1177/0741932507309716>
- Giancarlo-Gittens, C. A. 2010. Assessing critical thinking dispositions in an era of high-stakes standardized testing. In *Critical Thinking, Education and Assessment*, ed. J. Sobocan and L. Groake, 17–34. London, Ontario, Canada: The University of Western Ontario through the Althouse Press.
- Giancarlo, C. A., S. W. Blohm, and T. Urdan. 2004. Assessing secondary students' disposition toward critical thinking: Development of the California Measure of Mental Motivation. *Educational and Psychological Measurement* 64(2): 347–364. <http://dx.doi.org/10.1177/0013164403258464>
- Gittens, C. A., and P. A. Facione. 2014. *California Measure of Mental Motivation (CM3): An inventory of critical thinking dispositions. User Manual Supporting Levels IA, IB, II, and III Grades K-2, 3-5, 6-12, and Adults*. San Jose, CA: Insight Assessment / California Academic Press. <http://www.insightassessment.com/>
- Golding, J. 2001. The national numeracy strategy: Thoughts from a secondary classroom. *Mathematics in School* 30(1): 1–3.
- Graham, J. 2006. Congeneric and (essentially) tau-equivalent estimates of score reliability: What they are and how to use them. *Educational and Psychological Measurement* 66(6): 930–944. <http://dx.doi.org/10.1177/0013164406288165>
- Gray, E., M. Pinto, D. Pitta, and D. Tall. 1999. Knowledge construction and diverging thinking in elementary and advanced mathematics. *Educational Studies in Mathematics*, 38: 111–133. <http://dx.doi.org/10.1023/A:1003640204118>
- Halpern, D. F. 1996. *Thought and knowledge: An introduction to critical thinking* (3rd ed.). Mahwah, NJ: Lawrence Erlbaum.
- Hauser, R. M., C. F., Edley, J. A. Koenig, and S. W. Elliott. 2005. *Measuring literacy: Performance levels for adults*. Washington DC: National Academies Press. Retrieved from <http://www.nap.edu/openbook.php?isbn=0309096529>.
- Hoover, H. D., S. B. Dunbar, and D. A. Frisbie. 2001. *Iowa Test of Basic Skills – Form A*. Rolling Meadows, IL: Riverside Publishing.
- IA. See Insight Assessment, LLC.
- Insight Assessment, LLC. 2014. Retrieved from <http://www.insightassessment.com/>.
- Iowa Testing Programs. 2014. *ITBS Research Guide*. The University of Iowa College of Education. Retrieved from <http://itp.education.uiowa.edu/ia/ITBSResearchGuide.aspx>
- Kasof, J. 1997. Creativity and breadth of attention. *Creative Research Journal* 10: 303–315. http://dx.doi.org/10.1207/s15326934crj1004_2
- Kettler, T. 2014. Critical thinking among elementary school students: Comparing identified gifted and general education student performance. *Gifted Child Quarterly* 58(2): 127–136. <http://dx.doi.org/10.1177/0016986214522508>

- Kintsch, W. and J. G. Greeno. 1985. Understanding and solving word arithmetic problems. *Psychological Review* 92(1): 109–129.
<http://dx.doi.org/10.1037/0033-295X.92.1.109>
- Koedinger, K. R. and M. J. Nathan. 2004. The real story behind story problems: Effects of representations on quantitative reasoning. *Journal of the Learning Sciences* 13(2): 129–164. http://dx.doi.org/10.1207/s15327809jls1302_1
- Koedinger, K.R., and H.J.M. Tabachneck. 1994. Two strategies are better than one: multiple strategy use in word problem solving. Annual meeting of the American Educational Research Association, New Orleans, LA.
- Koedinger, K. R., M. W. Alibali, and M. J. Nathan. 1999. *A developmental model of algebra problem solving: Trade-offs between grounded and abstract representations*. Paper presented at the Annual Meeting of the American Educational Research Association (AERA). ERIC Doc No. ED433245.
- Komaroff, E. 1997. Effects of simultaneous violations of essential τ -equivalence and uncorrelated error of coefficient α . *Applied Psychological Measurement* 21: 337–348. <http://dx.doi.org/10.1177/01466216970214004>
- Kritzer, K. 2012. Building foundations for numeracy: A qualitative analysis of the basic concept knowledge demonstrated by young deaf children. *Australasian Journal of Early Childhood* 37(2): 106–112.
- Kuder G. F. and M. W. Richardson. 1937. The theory of the estimation of test reliability. *Psychometrika* 2: 151–160. <http://dx.doi.org/10.1007/BF02288391>
- Kuhn, D. 1990. Education for thinking: What can psychology contribute? In *Promoting cognitive growth over the lifespan*, ed. M. Schwebel, C. A. Mahler, and N. S. Fagley, 25–46. Hillsdale, NJ: Lawrence Erlbaum.
- Lipman, M. 1987. Some thoughts on the foundations of reflective education. In *Teaching thinking skills: Theory and practice*, ed. J. B. Baron and R. J. Sternberg, 151–161. New York: W. H. Freeman.
- Lohman, D. F. and E. P. Hagen. 2001. *Cognitive Abilities Test™ (CogAT®), Form 6*. Rolling Meadows, IL: Riverside Publishing.
- Lucas, N., S. Loo, and J. McDonald, 2005. Combining ‘subject knowledge’ with ‘how to teach’: an exploratory study of new initial teacher education for teachers of adult literacy, numeracy, and English for speakers of other languages. *International Journal of Lifelong Education* 24(4): 337–350.
<http://dx.doi.org/10.1080/02601370500169244>
- Mayer, R. E. 1997. Incorporating problem solving into secondary school curricula. In *Handbook of academic learning: Construction of knowledge—The educational psychology series*, ed. G. E. Phye, 473–492. San Diego, CA: Academic Press.
<http://dx.doi.org/10.1016/B978-012554255-5/50016-8>
- Mayes, R. L., J. H. Forrester, J. S. Christus, F. I. Peterson, R. Bonilla, and N. Yestness. 2014. Quantitative reasoning in environmental science: A learning progression. *International Journal of Science Education* 36(4): 635–658.
<http://dx.doi.org/10.1080/09500693.2013.819534>
- McNamara, O. and B. Corbin. 2001. Warranting practices: Teachers embedding the National Numeracy Strategy. *British Journal of Educational Studies*, 49(3): 260–284. <http://dx.doi.org/10.1111/1467-8527.t01-1-00175>

- Mendelsohn, G. A. 1976. Associative and attentional processes in creative performance. *Journal of Personality* 44: 341–369. <http://dx.doi.org/10.1111/j.1467-6494.1976.tb00127.x>
- Methe, S. A., J. M. Hintze, and R. G. Floyd. 2008. Validation and decision accuracy of early numeracy skill indicators. *School Psychology Review* 37(3): 359–373.
- Meyer, D. K., J. C. Turner, and C. A. Spencer. 1997. Challenge in a mathematics classroom: students' motivation and strategies in project-based learning. *Elementary School Journal* 97: 501–521. <http://dx.doi.org/10.1086/461878>
- Mononen, R., P. Aunio, and T. Koponen. 2014. A pilot study of the effects of RightStart instruction on early numeracy skills of children with specific language impairment. *Research in Developmental Disabilities* 35(5): 999–1014. <http://dx.doi.org/10.1016/j.ridd.2014.02.004>
- Moseley, C. and S. J. Ramsey. 2008. Elementary teachers' progressive understanding of inquiry through the process of reflection. *School Science & Mathematics* 108(2): 49–57. <http://dx.doi.org/10.1111/j.1949-8594.2008.tb17804.x>
- Myers, M. T. 2001. Preparing students for an uncertain future. *Liberal Education* 87(3): 22–25.
- Nathan, M. J. and K. R. Koedinger. 2000a. Teachers' and researchers' beliefs about the development of algebraic reasoning. *Journal for Research in Mathematics Education*, 31(2): 168–190. <http://dx.doi.org/10.2307/749750>
- . 2000b. An investigation of teachers' beliefs of students' algebra development. *Cognition and Instruction* 18(2): 209–237. http://dx.doi.org/10.1207/S1532690XCI1802_03
- National Council of Teachers of Mathematics. 2000. *Principles and standards for school mathematics*. Reston, VA: Author.
- . 2006. *Curriculum Focal Points for Prekindergarten through Grade 8 Mathematics: A Quest for Coherence*. Reston, VA: Author.
- . 2009. *Focus in High School Mathematics: Reasoning and Sense Making*. Reston, VA: Author.
- NCTM. See National Council of Teachers of Mathematics.
- Ojose, B. 2008. Applying Piaget's theory of cognitive development to mathematics instruction. *The Mathematics Educator* 18(1): 26–30.
- PARCC. See Partnership for Assessment of Readiness for College and Careers.
- Partnership for Assessment of Readiness for College and Careers. 2014. Retrieved from <http://www.parcconline.org/parcc-timeline>.
- Patrick, H., A. M. Ryan, and P. R. Pintrich. 1999. The differential impact of extrinsic and mastery goal orientations on males' and females' self-regulated learning. *Learning and Individual Differences* 11(2): 153–171. [http://dx.doi.org/10.1016/S1041-6080\(00\)80003-5](http://dx.doi.org/10.1016/S1041-6080(00)80003-5)
- Paul, R. 1984. Critical thinking: Foundation to education for a free society. *Educational Leadership* 42(15): 4.
- Peterson, P. L., S. R. Swing, K. D. Stark, and G. A. Waas. 1984. Students' cognitions and time on task during mathematics instruction. *American Educational Research Journal* 21(3): 487–515. <http://dx.doi.org/10.3102/00028312021003487>

- Pierce, R., H. Chick, and I. Gordon. 2013. Teachers' perceptions of the factors influencing their engagement with statistical reports on student achievement data. *Australian Journal of Education* 57(3): 237–255. <http://dx.doi.org/10.1177/0004944113496176>
- Pugalee, D. K., K. J. Hartman, and J. H. Forrester. 2008. Performance of middle grade students on quantitative tasks : A beginning dialogue on quantitative literacy in middle schools. *Investigations in Mathematics Learning* 1(2) : 35–51.
- Purpura, D. J. and C. J. Lonigan. 2013. Informal numeracy skills: The structure and relations among numbering, relations, and arithmetic operations in preschool. *American Educational Research Journal* 50(1): 178–209. <http://dx.doi.org/10.3102/0002831212465332>
- Quinnell, R., R. Thompson, and R. J. LeBard. 2013. It's not maths; it's science: Exploring thinking dispositions, learning thresholds, and mindfulness in science learning. *International Journal of Mathematical Education in Science and Technology* 44(6) : 808–816. <http://dx.doi.org/10.1080/0020739X.2013.800598>
- Riley, M.S., J. G. Greeno, and J. I. Heller. 1983. Development of children's problem-solving ability in arithmetic. In *The development of mathematical thinking*, ed. H. P. Ginsberg, 153–196. New York : Academic Press.
- Rivera-Batiz, F. L. 1992. Quantitative literacy and the likelihood of employment among young adults in the United States. *Journal of Human Resources* 27(2) : 313–328. <http://dx.doi.org/10.2307/145737>
- Riverside Publishing. 2010a. *Iowa Tests of Basic Skills® (ITBS®) Forms A, B, and C*. Retrieved from <http://www.riversidepublishing.com/products/itbs/index.html>.
- . 2010b. *The Cognitive Abilities Test™ (CogAT®), Form 6*. Retrieved from <http://www.riversidepublishing.com/products/cogAt/details.html#levels>.
- Root, R. 2009. Social justice through quantitative literacy: A course connecting numeracy, engaged citizenship and a just society. *Democracy & Education* 18(3): 37–43.
- Runco, M. A. and S. O. Sakamoto. 1999. Experimental studies of creativity. In *Handbook of creativity*, ed. R. J. Sternberg, 62–92. New York: Cambridge University Press.
- Seda Saracaloglu, A., H. Aktamis, and Y. Delioglu. 2011. The impact of the development of prospective teachers' critical thinking skills on scientific argumentation training and on their ability to construct an argument. *Journal of Baltic Science Education* 10(4): 243–260.
- Sezar, R. 2008. Integration of critical thinking skills into elementary school teacher education courses in mathematics. *Education* 128(3): 349–362.
- Simon, H. A. and S. K. Reed. 1976. Modeling strategy shifts in a problem-solving task. *Cognitive Psychology* 8: 86–97. [http://dx.doi.org/10.1016/0010-0285\(76\)90005-0](http://dx.doi.org/10.1016/0010-0285(76)90005-0)
- Skwarchuk, S., C. Sowinski, and J. LeFevre. 2014. Formal and informal home learning activities in relation to children's early numeracy and literacy skills: The development of a home numeracy model. *Journal of Experimental Child Psychology* 121: 63–84. <http://dx.doi.org/10.1016/j.jecp.2013.11.006>

- Smarter Balanced. See Smarter Balanced Assessment Consortium
 Smarter Balanced Assessment Consortium. 2014. Retrieved from
<http://www.smarterbalanced.org/smarter-balanced-assessments/>
- So, W. W. 2013. Connecting mathematics in primary science inquiry projects. *Journal of Science and Mathematics Education* 11: 385–406.
- Steen, L. A. 1990. Numeracy. In *Literacy in America*, special issue *Daedalus* 19(2): 211–231.
- . 1997. The new literacy. In *Why numbers count: Quantitative literacy for tomorrow's America*, ed. L. Steen, xv–xxviii. Princeton, NJ: College Board.
- . 1999. Numeracy: The new literacy for a data-drenched society. *Educational Leadership* 57(2): 8–13.
- . 2000a. Reading, writing and numeracy. *Liberal Education* 86(2): 26–37.
- . 2000b. The case for quantitative literacy. *Mathematics and Democracy: The Case for Quantitative Literacy*. Princeton, NJ: National Council on Education and the Disciplines and Educational Reform Initiative. Woodrow Wilson National Fellowship Foundation.
- . 2000c. *Mathematics, Numeracy, and Statistics*. Annual Congress of the Association of Mathematics Educators of South Africa (AMESA).
- Thompson, B. 2003. *Score Reliability: Contemporary Thinking on Reliability Issues*. Thousand Oaks, CA: SAGE.
- United States Department of Education. 1990. *National Goals for Education*. Washington, DC, U.S. Government Printing Office.
- van de Mortel, T. F., L. P. Whitehair, and P. M. Irwin. 2014. A whole-of-curriculum approach to improving nursing students' applied numeracy skills. *Nurse Education Today* 34(3): 462–467. <http://dx.doi.org/10.1016/j.nedt.2013.04.024>
- Vartanian, O. 2009. Variable attention facilitates creative problem solving. *Psychology of Aesthetics, Creativity, and the Arts* 3(1): 57–59. <http://dx.doi.org/10.1037/a0014781>
- , C. Martindale, and J. Kwiatkowski. 2007. Creative potential, attention, and speed of information processing. *Personality & Individual Differences* 34: 1370–1380.
- Ward, R.M., M. C. Schneider, and J. D. Kiper. 2011. Development of an assessment of quantitative literacy for Miami University. *Numeracy* 4(2): Article 4. <http://dx.doi.org/10.5038/1936-4660.4.2.4>
- Wells C. S. and J. A. Wollack. 2003. *An Instructor's Guide to Understanding Test Reliability*. Madison WI: University of Wisconsin Testing & Evaluation Services. Retrieved from <http://testing.wisc.edu/Reliability.pdf>.
- Wiest, L. R., H. J. Higgins, and J. H. Frost. 2007. Quantitative literacy for social justice. *Equity & Excellence in Education* 40(1): 47–55. <http://dx.doi.org/10.1080/10665680601079894>
- Wilkins, J. L. M. 2000. Preparing for the 21st century: The status of quantitative literacy in the United States. *School Science and Mathematics* 100(8): 405–418. <http://dx.doi.org/10.1111/j.1949-8594.2000.tb17329.x>
- Winn, I. J. 2004. The high cost of uncritical thinking. *Phi Delta Kappan* 85(7): 2. <http://dx.doi.org/10.1177/003172170408500706>

- Yang, Y. C. and C. Chang. 2013. Empowering students through digital game authorship: Enhancing concentration, critical thinking, and academic achievement. *Computers & Education* 68: 334–344. <http://dx.doi.org/10.1016/j.compedu.2013.05.023>
- Yang, D. and Wu W. 2010. The study of number sense: Realistic activities integrated into third-grade math classes in Taiwan. *The Journal of Educational Research* 103: 379–392. <http://dx.doi.org/10.1080/00220670903383010>
- Yee, L. and M. Simon. 2014. The role of health literacy and numeracy in contraception decision-making for urban Chicago women. *Journal of Community Health* 39(2): 394–399. <http://dx.doi.org/10.1007/s10900-013-9777-7>