Listening. Learning. Leading.®

Research Report
ETS RR–14-22

Assessing Quantitative Literacy in Higher Education: An Overview of Existing Research and Assessments With Recommendations for Next-Generation Assessment

Katrina Crotts Roohr

Edith Aurora Graf

Ou Lydia Liu

December 2014

ETS Research Report Series

Since its 1947 founding, ETS has conducted and disseminated scientific research to support its products and services, and to advance the measurement and education fields. In keeping with these goals, ETS is committed to making its research freely available to the professional community and to the general public. Published accounts of ETS research, including papers in the ETS Research Report series, undergo a formal peer-review process by ETS staff to ensure that they meet established scientific and professional standards. All such ETS-conducted peer reviews are in addition to any reviews that outside organizations may provide as part of their own publication processes. Peer review notwithstanding, the positions expressed in the ETS Research Report series and other published accounts of ETS research are those of the authors and not necessarily those of the Officers and Trustees of Educational Testing Service.

The Daniel Eignor Editorship is named in honor of Dr. Daniel R. Eignor, who from 2001 until 2011 served the Research and Development division as Editor for the ETS Research Report series. The Eignor Editorship has been created to recognize the pivotal leadership role that Dr. Eignor played in the research publication process at ETS.

ETS Research Report Series ISSN 2330-8516

RESEARCH REPORT

Assessing Quantitative Literacy in Higher Education: An Overview of Existing Research and Assessments With Recommendations for Next-Generation Assessment

Katrina Crotts Roohr, Edith Aurora Graf, & Ou Lydia Liu

Educational Testing Service, Princeton, NJ

Quantitative literacy has been recognized as an important skill in the higher education and workforce communities, focusing on problem solving, reasoning, and real-world application. As a result, there is a need by various stakeholders in higher education and workforce communities to evaluate whether college students receive sufficient training on quantitative skills throughout their postsecondary education. To determine the key aspects of quantitative literacy, the first part of this report provides a comprehensive review of the existing frameworks and definitions by national and international organizations, higher education institutions, and other key stakeholders. It also examines existing assessments and discusses challenges in assessing quantitative literacy. The second part of this report proposes an approach for developing a next-generation quantitative literacy assessment in higher education with an operational definition and key assessment considerations. This report has important implications for higher education institutions currently using or planning to develop or adopt assessments of quantitative literacy.

Keywords Quantitative literacy; quantitative reasoning; mathematics; numeracy; student learning outcomes; higher education; next-generation assessment

doi:10.1002/ets2.12024

Literacy is defined as “the ability to read and write” or “knowledge that relates to a specified subject” (Merriam-Webster, 2014, para. 1–2). Building from this definition, quantitative literacy has been defined as the ability to interpret and communicate numbers and mathematical information throughout everyday life (e.g., Organisation for Economic Co-Operation and Development [OECD], 2012b; Rhodes, 2010; Sons, 1996; Steen, 2001). Sharing many common characteristics with other related constructs, such as numeracy, quantitative reasoning, and mathematical literacy, quantitative literacy emphasizes skills related to problem solving, reasoning, and real-world application (Mayes, Peterson, & Bonilla, 2013; Steen, 2001). Unlike traditional mathematics and statistics, quantitative literacy is a “habit of mind” (Rhodes, 2010, p. 25; Steen, 2001, p. 5), focusing on certainty rather than uncertainty and data from the empirical world rather than the abstract (Steen, 2001, p. 5).

Quantitative literacy can be considered an essential element in society, especially in relation to many duties of citizens, such as the “allocation of public resources, understanding media information, serving on juries, participating in community organizations, and electing public leaders” (Steen, 2004, p. 28). The importance of quantitative literacy in society has been recognized by the higher education community (Rhodes, 2010). For instance, 91% of the member institutions of the Association of American Colleges and Universities (AAC&U) identified quantitative reasoning as an important learning outcome (AAC&U, 2011). Employers have also recognized the need for quantitative skills, insisting that all college graduates have quantitative skills regardless of their intended career path (National Survey of Student Engagement [NSSE], 2013a). In a recent online survey conducted by Hart Research Associates (2013), among the 318 employers surveyed about necessary skills for a successful college graduate in today’s economy, 90% stated that higher education institutions should continue to emphasize or increase the emphasis on a students’ ability to work with numbers and understand statistics (Hart Research Associates, 2013). Similarly, Casner-Lotto and Barrington (2006) found that among 400 surveyed employers, 64.2% identified mathematics as a very important basic knowledge skill for 4-year college graduates to be successful in today’s workforce. The authors also noted that basic mathematical skills underpin applied skills such as critical thinking and problem solving.

Corresponding author: K. C. Roohr, E-mail: kroohr@ets.org

Although the importance of quantitative literacy is recognized both in higher education and the workforce, many students do not feel prepared to use quantitative reasoning skills in the workplace. A survey conducted by McKinsey and Company (2013) was administered to 4,900 former Chegg (a textbook rental company) customers, which included a mix of 2- and 4-year college students graduating between 2009 and 2012. Among the students surveyed, 24% of 4-year college students and 34% of 2-year college students felt underprepared to use quantitative reasoning skills upon graduating college (McKinsey & Company, 2013). The underpreparedness of 2- and 4-year college students may be linked to the lack of student engagement in quantitative reasoning tasks in either a student's freshman year or student's senior year of college. For instance, the 2013 NSSE found that 49–63% of freshman (NSSE, 2013b) and 46–56% of senior (NSSE, 2013c) students either never or only sometimes reached conclusions based on their own analysis of numerical information, used numerical information to examine real-world problems, or evaluated other people's conclusions from numerical information. Results also found that students in fields other than science, technology, engineering, and mathematics (STEM; e.g., social science, education, communication, arts, and humanities) engaged in quantitative activities less often than their peers in STEM majors (NSSE, 2013a). Given the mismatch between college students' preparedness in quantitative literacy and the demands from stakeholders, there is an urgent need by various stakeholders in higher education and workforce communities to evaluate whether students receive sufficient training in quantitative skills in college.

Results from the Program for the International Assessment for Adult Competencies (PIAAC) also showed the underpreparedness of students' quantitative skills. PIAAC Numeracy measures adults' mathematical skills in real-world contexts. When focusing on adults aged 16 to 65 with bachelor's degrees, results showed that only 18% of US adults with a bachelor's degree scored in the top two proficiency levels (out of five) on the Numeracy measure, which was below an international average of 24% (Goodman et al., 2013). These results point to the critical need to understand why adult Americans are behind in quantitative literacy skills. Actions should be taken to delineate the various components underlying quantitative literacy, and quality assessments should be developed to identify students' strengths and weaknesses in quantitative literacy when they enter college.

The purposes of this report are to review and synthesize existing frameworks, definitions, and assessments of quantitative literacy, quantitative reasoning, numeracy, or mathematics and to propose an approach for developing a next-generation quantitative literacy assessment. We first examine how quantitative literacy is defined throughout the literature by various stakeholders with a focus in higher education. We then review existing assessments of quantitative literacy, quantitative reasoning, numeracy, or mathematics, considering both the structural and psychometric quality of those assessments. Following this review, we discuss challenges and issues surrounding the design of a quantitative literacy assessment. After reviewing and synthesizing the existing frameworks, definitions, and assessments, we propose an approach for developing a next-generation quantitative literacy assessment with an operational definition, framework, item formats, and task types. The goal of this article is to provide an operational framework for assessing quantitative literacy in higher education while also providing useful information for institutions developing in-house assessments. The next-generation assessment development should involve collaboration between institutions and testing organizations to ensure that the assessment has instructional value and meets technical standards.

Existing Frameworks, Definitions, and Assessments of Quantitative Literacy

Existing Frameworks and Definitions

Various terms have been used to represent the use of quantitative skills in everyday life, such as quantitative literacy, quantitative reasoning, numeracy, mathematical literacy, and mathematics (Mayes et al., 2013; Steen, 2001). These various terms have subtle differences in their definitions (Steen, 2001). Vacher (2014) attempted to decipher these subtle differences using WordNet, an online lexical database for English, and also found that the terms *numeracy*, *quantitative literacy*, and *quantitative reasoning* have subtle differences in their meaning, even though they are commonly treated as synonymous terms. Using WordNet, Vacher proposed four core components that correspond to these terms including: (a) "skill with numbers and mathematics," (b) "ability to read, write and understand material that includes quantitative information," (c) "coherent and logical thinking involving quantitative information," and (d) "disposition to engage rather than avoid quantitative information" (p. 11). The author proposed that numeracy includes (a), (b), and (d); quantitative literacy includes (b), (c), and (d); and quantitative reasoning includes (c) and (d) (Vacher, 2014). Note that these categorizations are also arbitrary.

With various terms being used, there has been some disagreement among faculty in higher education institutions about how quantitative literacy is defined (Steen, 2004). Despite this disagreement, definitions of quantitative literacy and similar constructs throughout the literature have many commonalities, as shown in Vacher (2014). Recognizing these commonalties is critical to develop a concrete definition of quantitative literacy. Definitions throughout the literature have been developed either for understanding what it means to be quantitatively literate or for developing assessments and curricula. This section describes frameworks and definitions of quantitative literacy and synonymous terms or constructs (e.g., quantitative reasoning, numeracy) identified in the literature by national and international organizations, workforce initiatives, higher education institutions and researchers, and K–12 theorists and practitioners.

Frameworks by National and International Organizations

AAC&U's Liberal Education and America's Promise (LEAP) and Lumina's Degree Qualifications Profile (DQP) are two higher education initiatives developed by national organizations that identify quantitative skills as an element of their frameworks. The LEAP initiative was launched in 2005 and emphasizes the importance of a 21st century liberal education (AAC&U, 2011). Similarly, the DQP tool was developed with the intent of transforming US higher education by clearly identifying what students should be expected to know and do upon earning an associate's, bachelor's, or master's degree (Adelman, Ewell, Gaston, & Schneider, 2011). Both initiatives discuss important educational outcomes at the college level, with LEAP focusing on outcomes for every college student (AAC&U, 2011) and DQP focusing on outcomes for college students at specific degree levels, regardless of student major (Adelman et al., 2011). As part of the LEAP initiative, a set of Valid Assessment of Learning in Undergraduate Education (VALUE) rubrics was developed for each learning outcome, including quantitative literacy. In defining quantitative literacy, both quantitative reasoning and numeracy are recognized as synonymous terms to quantitative literacy (Rhodes, 2010). The rubric identified six important skills of quantitative literacy: interpretation, representation, calculation, application/analysis, assumptions, and communication, each defined in terms of proficiency level (Rhodes, 2010). Alternatively, the DQP uses the term *quantitative fluency* and breaks down quantitative fluency into different categories based on degree level, discussing different skills such as interpretation, explanation of calculations, creation of graphs, translation of problems, construction of mathematical arguments, reasoning, and presentation of results in various formats (Adelman et al., 2014).

Similar efforts in defining quantitative literacy have been made by the American Mathematical Association of Two-Year Colleges (AMATYC; Cohen, 1995), the Mathematical Association of America (MAA; Sons, 1996), and the OECD (2012b). The AMATYC developed a clear set of standards for introductory college mathematics intended for college students obtaining either an associate's or a bachelor's degree, similar to the DQP. However, instead of describing various quantitative skills for students across degree levels, a framework for mathematics standards was developed, focusing on students' intellectual development, instructors' pedagogical practices, and curricular content in higher education. The OECD (2012a) developed a framework with four facets of numeracy—contexts, responses, mathematical content/information/ideas, and representations—as well as a list of enabling factors and processes, such as the integration of mathematical knowledge and conceptual understanding of broader reasoning, problem-solving skills, and literacy skills. Alternatively, the MAA simply provided a list of five skills that every college student should have to be quantitatively literate, emphasizing skills such as interpretation, representation, problem solving, and estimation (Sons, 1996). These mathematical skills are similar to those enumerated by other national and international associations. Quantitative literacy definitions from these national and international organizations can be found in Table 1.

Frameworks by Workforce Initiatives

The US federal government and workforce initiatives have also recognized the importance of student learning outcomes but have focused on the term mathematics. The Employment and Training Administration's Industry Competency Model, developed by the US Department of Labor (USDOL), models essential skills and competencies for the workplace, specifically, economically important industries in the health, technology, and science fields (USDOL, 2013). This model, unlike the models developed by national and international organizations, is represented by stacked building blocks with more general competencies at the bottom building to more narrow competencies at the top. The second block from the bottom, the academic block, defines mathematics in terms of quantification, computation, measurement and estimation, and

Table 1 Definitions of Quantitative Literacy From National and International Organizations

Framework	Definition
AAC&U's Liberal Education and America's Promise	"Quantitative literacy (QL) — also known as numeracy or quantitative reasoning — is a 'habit of mind,' competency, and comfort in working with numerical data. Individuals with strong QL skills posses the ability to reason and solve quantitative problems from a wide array of authentic contexts and everyday life situations. They understand and can create sophisticated arguments supported by quantitative evidence and they can clearly communicate those arguments in a variety of formats (using words, tables, graphs, mathematical equations, etc., as appropriate)" (Rhodes, 2010, p. 25).
Lumina's Degree Qualifications Profile 2.0	"The student [at the bachelor's level] translates verbal problems into mathematical algorithms as to construct valid arguments using the accepted symbolic system of mathematical reasoning and presents the resulting calculations, estimates, risk analyses or quantitative evaluations of public information in papers, projects or multimedia presentations. The student constructs mathematical expressions for complex issues most often described in non-quantitative terms" (Adelman et al., 2014, p. 22).
Mathematical Association of America (MAA)	"A college student who is considered quantitatively literate should be able to: 1. Interpret mathematical models such as formulas, graphs, tables, and schematics, and draw inferences from them. 2. Represent mathematical information symbolically, visually, numerically, and verbally. 3. Use arithmetical, algebraic, geometric and statistical methods to solve problems. 4. Estimate and check answers to mathematical problems in order to determine reasonableness, identify alternatives, and select optimal results. 5. Recognize that mathematical and statistical methods have limits" (Sons, 1996, Part II, para. 6).
Organisation for Economic Co-Operation and Development (OECD)	"The ability to access, use, interpret and communicate mathematical information and ideas in order to engage in and manage the mathematical demands of a range of situations in adult life. To this end, numeracy involves managing a situation or solving a problem in a real context, by responding to mathematical content/information/ideas represented in multiple ways" (OECD, 2012b, p. 20).

application, defining important content within each skill area (USDOL, 2013). Another workforce-based definition for mathematics was developed by Capital Workforce Partners (2014), with a list of career competency standards based on a range of interviewed employees. These standards include quantitative skills such as a person's basic ability to do mathematics, apply mathematics to business, create tables and graphs, integrate information, and use mathematical functions (Capital Workforce Partners, 2014).

Frameworks by Higher Education Institutions and Researchers

In addition to the higher education initiatives by the AAC&U and Lumina Foundation, institutions have developed in-house frameworks that guide quantitative literacy or quantitative reasoning assessments and coursework. Many of these in-house frameworks are similar in structure to the AAC&U VALUE rubric with a list of skills at different quantitative literacy proficiency levels (e.g., Samford University, 2009; University of Kentucky, 2012). Other institutions, such as Michigan State University, have developed standards for students at three different stages of quantitative literacy development, a similar approach to Lumina's DQP (Estry & Ferrini-Mundy, 2005). Alternatively, like the MAA, some institutions simply list the skills required of a quantitatively literate individual (e.g., Mount St. Mary's College, 2013). Compared with the large-scale higher education frameworks and those of national organizations such as the MAA, the definitions of quantitative literacy show considerable overlap, including skills such as application, evaluation of arguments, quantitative expression, interpretation, reasoning, and problem solving.

Like higher education institutions, researchers have also attempted to construct definitions, frameworks, and standards for quantitative literacy. For example, Steen (2001) identified 10 quantitative literacy elements such as confidence with mathematics, interpreting data, logical thinking, mathematics in context, and number and symbol sense. Likewise, Mayes et al. (2013) developed a framework for quantitative reasoning in the context of science, focusing on components such as quantification act (i.e., identifying objects, observing attributes, and assigning measures), and quantitative literacy, interpretation, and modeling. Among the definitions developed by researchers, many have defined quantitative literacy in terms of application to real-world problems (Hollins University, 2013; Kirsch, Jungeblut, Jenkins, & Kolstad, 2002; National Numeracy Network [NNN], 2013; OECD, 2000; Ward, Schneider, & Kiper, 2011), or in terms of reasoning skills (J. Bennett & Briggs, 2008; Hollins University, 2013; Langkamp & Hull, 2007; NNN, 2013; Steen, 2004).

Frameworks and Standards by K–12 Experts and Practitioners

The most well-known K–12 standards relevant to quantitative literacy are the Common Core State Standards for Mathematics developed by the Council of Chief State Officers (CCSSO) and the National Governors Association (NGA) for Best Practices. Although developed for K–12 with a focus on standards for mathematics that should be taught in school, the Common Core State Standards for Mathematics were constructed to help improve students' college and career readiness in terms of quantitative knowledge and skills, identifying specific mathematical content areas and competencies students need to master, such as problem solving, reasoning, modeling, and expression, within the content areas of number and quantity, algebra, functions, modeling, geometry, and statistics and probability (NGA & CCSSO, 2010). These various skills identified in the Common Core State Standards for Mathematics are highly related to many of the higher education and workforce definitions of quantitative literacy and quantitative reasoning.

The American Diploma Project (Achieve, Inc., 2004) also linked K–12 education to postsecondary education and careers. This project established a set of English and mathematical skills and benchmarks that high school graduates should master to be successful in their future endeavors. Mathematics benchmarks were organized into four content strands: (a) number sense and numerical operations, (b) algebra, (c) geometry, and (d) data interpretation, statistics, and probability. The American Diploma Project also noted that mathematical skills are crosscutting and involve a student's ability to blend knowledge and skills when problem solving, to connect new information with existing knowledge, and to access and assess knowledge from a variety of sources (Achieve, Inc., 2004), which are common skills identified within quantitative literacy. These mathematical skills and benchmarks in both the Common Core State Standards and American Diploma Project are comparable to many of the skills identified within higher education and workforce initiatives.

Another set of K–12 frameworks, focusing more on noncognitive skills within core subject areas, includes the Partnership for 21st Century Skills (P21) Math Map (Saltrick et al., 2011). This framework differs from other frameworks by focusing on mathematical content knowledge and mathematical processes integrated with 21st century skills such as creativity and innovation, critical thinking and problem solving, communication and collaboration, and other noncognitive skills. This framework is intended to make teaching and learning of mathematics more engaging, relevant, and rigorous for students (Saltrick et al., 2011).

Existing Assessments Measuring Quantitative Literacy Skills

A number of tests and subtests assess the quantitative literacy, quantitative reasoning, numeracy, or mathematics skills of students in higher education. Most of these assessments are multiple-choice tests administered on a computer. Table 2 summarizes these existing college-level and adult-level assessments, which include the three assessments approved by the Voluntary System of Accountability Program (VSA) to provide evidence of student learning outcomes in colleges and universities: the Collegiate Assessment of Academic Proficiency (CAAP), Collegiate Learning Assessment+ (CLA+), and the *ETS*® Proficiency Profile (EPP; VSA, 2013). Other assessments measuring quantitative literacy, quantitative reasoning, numeracy, or mathematics include the College-Level Examination Program (*CLEP*®), Graduate Management Admissions Test (GMAT), the *GRE*® General Test, National Assessment of Adult Literacy (NAAL), PIAAC, and two assessments developed by Insight Assessment, including Quant Q and the Test of Everyday Reasoning—Numeracy (TER-N).

In addition to these widely used assessments measuring quantitative literacy skills, many institutions have developed their own quantitative assessments. For example, the University of Cambridge developed an essay assessment called the Sixth Term Examination Papers in Mathematics (STEP) to evaluate pure mathematics, mechanics, and probability and

Table 2 Existing Assessments Measuring Quantitative Literacy Skills

Test	Developer	Format	Delivery	Length	# Items	Themes/topics
College-Level Examination Program (CLEP) College Mathematics	College Board	Multiple choice; multiple-selection multiple choice; numeric entry	Computer	90 min	60 items (not all items contribute to final score)	Measures the examinees' ability to solve routine, straightforward problems, and nonroutine problems that require an understanding of concepts and application of skills and concepts. Topics on the assessment include sets, logic, real number system, functions and their graphs, probability and statistics, algebra, and geometry (College Board, 2012).
Collegiate Assessment of Academic Proficiency (CAAP) Mathematics	ACT	Multiple choice	Paper/pencil	40 min	35 items	Assesses proficiency in solving mathematical problems encountered in many postsecondary curricula, emphasizing quantitative reasoning rather than memorization of formulas. The content areas tested include (a) pre-algebra, (b) elementary, intermediate, and advanced algebra, (c) coordinate geometry, and (d) trigonometry (CAAP Program Management, 2012).
Collegiate Learning Assessment+ (CLA+) Scientific and Quantitative Reasoning (SQR)	Council for Aid to Education (CAE)	Multiple choice	Computer	30 min (not a distinct subtest—SQR items are within the 30 min period)	10 SQR items (out of 26 total items on the full CLA+)	A set of 10 multiple-choice items all attached to documents that emulate real-world scenarios or problems in a work or academic environment. Documents include reference sources such as data tables or graphs, a newspaper article, research report, etc. These multiple-choice items require careful analysis and evaluation of information by examinees (Zahner, 2013).
ETS Proficiency Profile (EPP) Mathematics	Educational Testing Service (ETS)	Multiple choice	Computer and paper/pencil	Approximately 30 min (full test is 2 hours)	27 items (standard form)	Assesses the ability to "recognize and interpret mathematical terms; read and interpret tables and graphs; evaluate formulas; order and compare large and small numbers; interpret ratios, proportions, and percentages; read scientific measuring instruments; recognize and use equivalent mathematical formulas or expressions" (ETS, 2010, p. 4).

Table 2 Continued

Test	Developer	Format	Delivery	Length	# Items	Themes/topics
Graduate Management Admissions Test (GMAT) Quantitative	Graduate Management Admission Council (GMAC)	Multiple choice	Computer	75 min	37 items	Measures the ability to reason quantitatively, solve quantitative problems, and interpret graphical data. Both problem-solving and data-sufficiency questions are used and require the knowledge of arithmetic, elementary algebra, and commonly known concepts of geometry (GMAC, 2013a).
Graduate Record Examinations® (GRE) Quantitative Reasoning Measure	ETS	Multiple choice; multiple-selection multiple choice; numeric entry	Computer and paper/pencil	70 min	40 items	Measures the ability to interpret and analyze quantitative information and use mathematical skills in arithmetic, algebra, geometry, and data interpretation to solve problems (ETS, 2013b).
National Assessment of Adult Literacy (NAAL) Quantitative Literacy	US Department of Education	Open-ended/short answer	Paper/pencil	Untimed	47 items	Assesses the ability to "identify, describe, or perform an arithmetic operations (addition, multiplication, subtraction, and division) either in prose or document materials" (Institute of Educational Statistics, n.d., para. 5).
Programme for the International Assessment of Adult Competencies (PIAAC) Numeracy	Organisation for Economic Co-Operation and Development (OECD)	Multiple choice; clicking/selecting objects; numeric entry; highlighting objects	Computer and paper/pencil	Around 60 min (but is not timed)	56 items	Measures the ability to solve problems in real contexts (everyday life, work, society, further learning) by responding (identify, locate or access; act upon and use: order, count, estimate, compute, measure, model; interpret; evaluate/analyze; communicate) to mathematical content (quantity and number; dimension and shape; pattern, relationships and change; data and chance), represented in multiple ways (objects and pictures; numbers and symbols; formulae; diagrams and maps, graphs, tables; texts; technology-based displays) (OECD, 2012b).

Table 2 Continued

Test	Developer	Format	Delivery	Length	# Items	Themes/topics
Quant Q	Insight assessment	Multiple choice	Computer and paper/pencil	50 min	28 items	Assesses basic mathematical knowledge and integration of critical thinking skills, as well as quantitative reasoning. Score reports indicate the test taker's skills in pattern recognition, probability combinatorics, geometry and optimization, and out-of-the-box algebra (i.e., items involving algebra or other more basic mathematical techniques) (Insight Assessment, 2013a).
Test of Everyday Reasoning—Numeracy (TER-N)	Insight assessment	Multiple choice	Computer and paper/pencil	Not available	40 items	Measures quantitative reasoning in addition to critical thinking skills. Score reports indicate the test taker's skills in numeracy and reasoning skills, including overall general reasoning, analysis, interpretation, evaluation, inference, explanation, induction, and deduction (Insight Assessment, 2013b).

statistics (Admissions Testing Service, 2013). The STEP Mathematics is also used across other institutions in the United Kingdom such as the University of Warwick (University of Warwick, 2013). Similarly, the Center for Assessment and Research Studies at James Madison University developed a multiple-choice assessment called the Quantitative Reasoning Test, Version 9 (QR-9). This assessment measures learning objectives such as the use of different mathematical methods to analyze, organize, and interpret different phenomena. The assessment also evaluates a student's ability to discriminate between association and causation (Sundre, 2008). Other measures developed to assess quantitative literacy or quantitative reasoning include the Quantitative Literacy Skills Assessment by Colby-Sawyer College (Steele & Kilic-Bahi, 2008), the Quantitative Literacy Assessment by Miami University (Ward et al., 2011), the Quantitative Reasoning Assessment by Wellesley College (Wellesley College, n.d.), and Carleton College's Quantitative Inquiry, Reasoning, and Knowledge (QuIRK) initiative (Carleton College, 2013).

Similarly, a number of K–12 assessments target aspects of quantitative literacy such as the Programme for International Student Assessment (PISA) Mathematics, an international assessment measuring mathematical literacy for 15-year-old students, and PISA Financial Literacy, an international assessment for 15-year-olds measuring student knowledge and application of both financial concepts and risks (OECD, 2013). In addition to international assessments, national K–12 accountability mathematics assessments have been built using the Common Core State Standards, such as the Partnership for Assessment of Readiness for College and Careers (PARCC, 2014), and the Smarter Balanced Assessment Consortium (SBAC, n.d.). Likewise, a research and development initiative is being conducted at the Educational Testing Service (ETS) on a K–12 accountability measure called the Cognitively-Based Assessment *of*, *for*, and *as* Learning (*CBAL*™), with one of the content areas being mathematics. The goal of this initiative is to unify three main components: accountability, formative assessment, and professional support (ETS, 2014a).

In the following sections we discuss the test content, contexts, item types, calculator use, test reliability, and validity evidence, including convergent, concurrent, and predictive validity evidence.

Test Content and Contexts

The existing assessments measuring quantitative literacy skills assess a variety of content areas and contexts. Content is defined as the mathematical knowledge and skills needed to answer a question, and context is defined as the setting described in the question (Dwyer, Gallagher, Levin, & Morley, 2003). The assessed content is identified for all assessments except the CLA+ Scientific and Quantitative Reasoning (SQR), with the most commonly identified content consisting of geometry and measurement, algebra, probability and statistics, number sense, arithmetic, and pre-algebra (see Table 3). Additionally, items across assessments are written both to pure mathematical contexts and to applied contexts. Pure mathematical contexts include items that assess strict mathematical content such as solving an algebraic expression. Existing assessments with a proportion of test items written to a pure mathematical context include CAAP Mathematics, CLEP Mathematics, EPP Mathematics, GMAT Quantitative, and the GRE Quantitative Reasoning measure. Applied contexts vary across assessments and include contexts such as real-world scenarios (GRE Quantitative Reasoning), accompanying documentation (e.g., newspaper articles, data tables, emails; CLA+ SQR), problems encountered in postsecondary curricula (CAAP Mathematics), and specific disciplines (e.g., humanities, social sciences, and natural sciences; EPP Mathematics). PIAAC Numeracy has the most clearly defined contexts, with items written to work-related, personal, society and community, and education and training contexts (OECD, 2012b).

Item Format

Single- and multiple-selection multiple-choice items are the most commonly used item formats throughout the 10 assessments measuring quantitative literacy skills. For a single-selection multiple-choice item, the answer key consists of only one correct choice, while for a multiple-selection multiple-choice item, the answer key consists of one or more choices that satisfy the conditions specified in the question. Single-selection multiple-choice items are used across all of the assessments, and multiple-selection multiple-choice items are used by both the GRE Quantitative Reasoning and the PIAAC Numeracy. Some assessments (e.g., CLA+ SQR) also group multiple-choice items together using a common stimulus such as a table, graph, or other data presentation.

Table 3 Test Content on Existing Assessments Measuring Quantitative Literacy Skills

	Geometry and measurement	Algebra	Probability and statistics	Number sense	Arithmetic	Pre-algebra	Pattern recognition	Trigonometry
Collegiate Assessment of Academic Proficiency (CAAP) Mathematics	X	X				X		X
College-Level Examination Program (CLEP) Mathematics	X	X	X	X				
ETS Proficiency Profile (EPP) Mathematics	X	X	X	X	X	X		
Graduate Management Admissions Test (GMAT) Quantitative	X	X	X	X	X	X		
GRE Quantitative Reasoning	X	X	X	X	X	X		
National Assessment of Adult Literacy (NAAL) Quantitative Literacy					X			
Program for the International Assessment for Adult Competencies (PIAAC) Numeracy	X	X	X	X			X	
Quant Q	X	X	X				X	
Test of Everyday Reasoning—Numeracy (TER-N)					X			
Total	7	7	6	5	5	4	2	1

Multiple-choice format lends itself to various task types such as the quantitative comparison task found in the GRE Quantitative Reasoning measure and the data-sufficiency task in the GMAT Quantitative section. The quantitative comparison task involves the comparison of two quantities and asks the examinee to determine whether one of the quantities is greater than, less than, or equal to the other quantity, or whether the relationship is indeterminable based on the information provided (ETS, 2013c). The data-sufficiency task involves two statements and asks whether the statements provide sufficient information to answer a given question (GMAC, 2013a).

Computer delivery of an assessment allows for variations on traditional multiple-choice response formats, such as items involving the clicking or highlighting of objects, which are used by PIAAC Numeracy. For instance, for a multiple-selection multiple-choice item, instead of selecting multiple checkboxes, an examinee could click on multiple bars on a bar graph. In addition to multiple-choice items, another common item format across assessments is numeric entry, where an examinee enters a numeric value as the response rather than selecting one from a list of choices. Numeric entry is used by CLEP Mathematics, GRE Quantitative Reasoning, and PIAAC Numeracy.

Calculator Use

An important consideration with any assessment of mathematics is whether a calculator will be permitted. Existing higher education quantitative assessments such as the EPP and CAAP allow calculators but stress that all problems can easily be solved without a calculator. The PIAAC Numeracy measure also permits calculator use, recognizing that calculators are easily available when conducting quantitative tasks throughout everyday life (PIAAC Numeracy Expert Group, 2009). Similarly, the GRE Quantitative Reasoning measure allows an examinee to use a calculator to help shorten the time it takes to perform computation; however, it is noted that the calculator is provided solely as a supplement to the assessment and does not replace the examinee's knowledge of mathematics (ETS, 2014b).

The use of a calculator can be advantageous for a quantitative literacy assessment and can improve construct validity by allowing the examinee to focus on problem-solving skills rather than strict computation of a test item (Bridgeman, Harvey, & Braswell, 1995). Calculator use has also been found to improve mathematical problem-solving strategies and positively influence students' attitudes toward mathematics (Ellington, 2003). It is important to think about the impact of having a calculator on a quantitative literacy assessment. Although Bridgeman et al. (1995) found that construct validity can be improved with the use of a calculator, the authors also found that in some cases construct validity can decrease. A major advantage to having a computer-based assessment is that developers can easily include some items that allow a calculator and some items that do not allow a calculator (e.g., questions on estimation) while also controlling the calculator features (e.g., basic vs. scientific vs. graphing).

Test Reliability

Reliability refers to the consistency of measures (American Educational Research Association [AERA], American Psychological Association, & National Council on Measurement in Education, 1999; Traub & Rowley, 1991). Methods for estimating reliability include parallel form, test–retest, split-half, internal consistency, and interrater reliability. Both parallel form and test–retest reliability estimates require multiple test administrations, whereas split-half and internal consistency (e.g., coefficient α) estimates are derived from items within a single test administration. To estimate reliability on human-scored constructed-response items, interrater reliability is estimated by calculating the score agreement across multiple raters. Test length is highly related to reliability, with tests with a larger number of items typically yielding higher reliability estimates than tests with a smaller number of items (Traub & Rowley, 1991). For the same reason, a multiple-choice test typically has higher reliability than a constructed-response test, as more multiple-choice items can be administered than constructed-response items within the same time frame.

As discussed previously, many of the existing tests measuring quantitative literacy skills use multiple-choice items and have published results on test or subscale reliability. For instance, satisfactory reliability estimates have been found on the EPP Mathematics standard form with estimates around .85 (ETS, 2010; Lakin, Elliott, & Liu, 2012), on the CLEP College Mathematics with estimates around .90 (College Board, 2012), and on the CAAP Mathematics with estimates of .95 and .93 for freshman and senior students, respectively (Klein et al., 2009). Satisfactory reliability estimates have also been found for both the GRE Quantitative Reasoning measure and GMAT Quantitative section with reliability estimates of .95 and .90, respectively (ETS, 2013a; GMAC, 2013b). For both assessments with constructed-response items (i.e., NAAL Quantitative Literacy and PIAAC Numeracy), no information was found on the internal consistency of those measures; however, information on interrater reliability was reported. For example, the 2003 NAAL Quantitative Literacy showed high percent agreement between raters ranging from 92.6% to 100% (Baldi, 2009), and PIAAC Numeracy's high percent agreement was 99.1% within-country and 96.7% across countries (Tamassia, Lennon, & Yamamoto, 2013).

Convergent Validity Evidence

Convergent validity evidence looks at the relationship between scores across tests measuring similar constructs (AERA et al., 1999). Klein et al. (2009) examined the relationship among the three approved VSA measures and found a strong relationship between EPP and CAAP Mathematics with a student-level correlation of .76 and a school-level correlation of .98. These results provide evidence that both the EPP and CAAP Mathematics sections are measuring a similar construct. At the time of this study, the CLA did not have an equivalent quantitative literacy section to examine this relationship.

Concurrent Validity Evidence

Concurrent validity refers to the relationship between a predictor and a criterion measured at the same time rather than at a later time (AERA et al., 1999). Concurrent validity has been evaluated for EPP Mathematics by examining the relationship between student performance on EPP Mathematics and grade point average (GPA), finding that across a 10-year period, students with higher GPA consistently yielded higher EPP Mathematics scores (Liu & Roohr, 2013). A similar relationship was found between test takers' EPP Mathematics scores and the number of college credit hours they had taken (Liu & Roohr, 2013). These results suggest that indicators of students' success, such as GPA and the number of credit hours completed, are strongly associated with the level of performance on EPP Mathematics.

Predictive Validity Evidence

Predictive validity refers to how well particular outcomes of interest measured at a later time (e.g., first-year graduate student GPA) are predicted from test scores on an assessment that purports to measure relevant constructs (e.g., GRE;

AERA et al., 1999). Although some of the existing assessments (i.e., CAAP, CLA+, EPP) measure certain aspects of college learning outcomes, results from these assessments may also predict other college-level outcomes. To date, predictive validity for the assessments measuring quantitative literacy skills has been examined by looking at a variety of school-level outcomes. For example, moderate correlations ranging from .23 to .48 have been found between GPA or grades and test scores on CAAP Mathematics, GMAT Quantitative, and GRE Quantitative Reasoning (CAAP Program Management, 2012; GMAC, 2013b; Kuncel, Hezlett, & Ones, 2001). Small to moderate correlations have also been found between EPP Mathematics scores and credit hours or courses completed (Lakin et al., 2012; Marr, 1995). Other investigated school-level outcomes have included faculty ratings, comprehensive exam scores, publication citation count, degree attainments, time to complete, and research productivity. Operational predictive validity evidence (i.e., correlations with corrections for range restriction and criterion unreliability) has ranged from .11 to .47 between these additional school-level outcomes and GRE Quantitative Reasoning test scores (Kuncel et al., 2001).

It is evident that much of the existing predictive validity evidence has focused on the prediction of college-level outcomes; however, more predictive validity evidence is needed after students leave college. Essentially, future research should consider using a next-generation quantitative literacy assessment to predict long-term life outcomes. For instance, future research should evaluate the relationship between the assessment scores and whether a student can make sound quantitative decisions in life, such as making a decision between renting or buying a property. Making sound financial decisions was identified as a critical content area for college graduates in the workforce (Casner-Lotto & Barrington, 2006), so obtaining this evidence could help to predict whether students will have those skills related to financial decisions and other related quantitative skills that are common in the workforce.

Broad Issues in Assessing Quantitative Literacy in Higher Education

In developing a new assessment, it is important to consider challenges and broad issues in assessing that construct. Recognizing these challenges and issues during test development can help to ensure a reliable, valid, and fair assessment for examinees that is commensurate with the stakes of the assessment. This section describes a set of issues pertaining to the assessment of quantitative literacy in higher education.

Mathematics Versus Quantitative Literacy

When assessing quantitative literacy, it is important to understand the difference between quantitative literacy and traditional mathematics. Steen (2001) clearly addressed this difference, stating that mathematics typically focuses on a "Platonic realm of abstract structures," whereas quantitative literacy is more "anchored in data derived from and attached to the empirical world" (p. 5). Steen also noted that there is a difference between quantitative literacy and statistics. He stated that statistics is "primarily about uncertainty," whereas quantitative literacy is mainly about the "logic of certainty" (p. 5).

Quantitative literacy is distinctively different from mathematics and involves solving problems using mainly primary- and secondary-level mathematics within a particular context, such as the workplace, focusing on the student's ability to use reasoning skills to address those context-specific real-world problems (Steen, 2001, 2004). Another distinction between quantitative literacy and mathematics is that mathematics is typically practiced on its own as a discipline, whereas quantitative literacy is typically employed alongside other literacies (Ewell, 2001), such as reading and writing. The difference between quantitative literacy and mathematics can also be found across various assessments. For example, items on the *SAT*® and ACT are typically decontextualized and focus on strict mathematical content (Steen, 2004). Alternatively, quantitative literacy questions can be made very difficult using basic mathematical content and increasing the complexity of mathematical reasoning processes to reach a solution (Dwyer et al., 2003). Recognizing and understanding these differences between quantitative literacy and mathematics is critical when developing a quantitative literacy assessment.

General Versus Domain Specific

The quantitative skills a student is expected to master can vary based on the student's major, and this raises the following question: Should an assessment of quantitative literacy be domain specific or more general? There is no question that students pursuing a mathematics or science degree will take more quantitative courses than students pursuing an English or history degree. However, regardless of a student's major or career path, all students should receive context-rich quantitative literacy as part of their college education (Ewell, 2001; NSSE, 2013a) and should be able to "draw information from charts, graphs, and geometric figures," and have the "ability to complete straightforward estimations and calculations" (Rhodes, 2010, p. 25). In the 2005 study sponsored by the Conference Board of the Mathematical Sciences (CBMS2005; Lutzer, Rodi, Kirkman, & Maxwell, 2007), quantitative requirements were examined across 4-year institutions focusing on the requirements within a college at the institution (e.g., College of Arts and Sciences) with a mathematics department. Results indicated that 9 of 10 of these colleges at 4-year institutions had a quantitative requirement in Fall 2005 with courses such as calculus, elementary statistics, college algebra, pre-algebra, and special general education courses satisfying this requirement (Lutzer et al., 2007). These results show that all students, regardless of major, are being required to take quantitative courses.

Quantitative skills are not only an important part of a college education but are also needed in many aspects of society, such as dealing with money, working with schedules, or reading mainstream media. Developing an assessment to target more general quantitative literacy skills will allow for a broader representation of competencies, as well as those that are transferrable across majors. In the redesign of the DQP, Adelman et al. (2014) recognized the need for a framework to represent "broad, integrative studies and crosscutting proficiencies that graduates need for continued learning in complex and changing environments" (p. 12). The authors recognized that students change their jobs many times throughout their life, so having these general skills is important. Additionally, given that approximately 27% of students do not end up pursuing a career related to their major field of study (Abel & Deitz, 2013), it may be advantageous for students to take a more general quantitative literacy assessment rather than a domain-specific assessment.

Total Score Versus Subscores

Another challenge in developing a quantitative literacy assessment is whether to report subscores to examinees. The advantage of reporting subscores is that it provides score users with diagnostic information about their strengths and weaknesses; however, this advantage only holds if the subscores are sufficiently reliable and distinct from each other. As previously discussed, reliability is impacted by test length (Sinharay, Puhan, & Haberman, 2011; Traub & Rowley, 1991), and a short assessment is likely to yield lower reliability estimates. This phenomenon was found when examining the reliability of the CLA+ SQR section. With only 10 multiple-choice items, reliability estimates were quite low at .42 and .62 on Forms A and B, respectively (Zahner, 2013). These results suggest that caution needs to be used when reporting subscores based on a small number of test items.

It is not only important to consider the number of items within a subscore, but also to determine what diagnostic information the subscores should represent. A quantitative literacy assessment measures mathematical skills and content within varying real-world contexts. These skills and content areas are likely to overlap, making it difficult to distinguish which items belong to which specific subscore. Score users may also have different preferences for how the score information should be presented. For example, an employer may value a holistic score to evaluate a job candidate's overall quantitative competency, whereas a higher education institution may prefer subscores so actions may be taken to improve the areas that showed deficiencies. As a result, research should be conducted to determine the appropriateness of subscore reporting for a next-generation quantitative literacy assessment.

Students' Test-Taking Motivation

Students' motivation in taking low-stakes assessment is a major challenge to the validity of the test scores. In a recent study by Liu, Bridgeman, and Adler (2012), the authors found significant differences in performance among motivated and unmotivated students taking a low-stakes learning outcomes assessment in higher education. Without considering the role of motivation, invalid conclusions could be made about a higher education institution in relation to different learning outcomes. Liu et al. (2012) showed that nonfinancial incentives could successfully increase an individual student's motivation and therefore increase test scores. When developing a next-generation quantitative literacy assessment, a testing program or sponsor could consider offering a certificate of achievement to increase the relevance of the test to students so they may take it more seriously. This certificate could also be presented to future employers to show proficiency in quantitative literacy skills.

Proposing a Next-Generation Quantitative Literacy Assessment Framework

Using the information from the above review of definitions, frameworks, and assessments, we propose a next-generation quantitative literacy framework designed to measure the degree to which students in 2- and 4-year higher education institutions are able to identify and solve mathematical problems of practical importance. The framework is intended for all college students, regardless of major, and will inform the development of a quantitative literacy assessment that could be used for college accreditation, instructional improvement within the institution, and assessing individual students' proficiency levels.

Operational Definition

Development of an operational or construct definition is "essential to validity and fairness to the scientific integrity of the inferences drawn from assessments" and helps to aid "decision-making about assessment design, development, and interpretation" (Dwyer et al., 2003, p. 1). Based on a synthesized review of the frameworks and definitions by national and international organizations, workforce initiatives, higher educational institutions and researchers, and K–12 theorists and practitioners, as well as existing assessments measuring quantitative literacy skills, we propose an operational definition for a next-generation quantitative literacy assessment. The proposed definition is three-dimensional, addressing problem-solving skills, content, and contexts. Problem-solving skills are central to many of the frameworks and definitions of quantitative literacy, quantitative reasoning, and numeracy. Similarly, mathematical content is central to many of the assessments measuring quantitative literacy skills. Lastly, a clear distinction between mathematics and quantitative literacy is that quantitative literacy involves the application of primary and secondary-level mathematics to real-world problems, thus stressing the importance of contexts. As a result of our synthesized review, we define quantitative literacy as follows:

> Quantitative literacy is the comprehension of mathematical information in everyday life, and the ability to detect and solve mathematics problems in authentic contexts across a variety of mathematical content areas. Solving these applied mathematical problems includes (a) interpreting information, (b) strategically evaluating, inferring, and reasoning, (c) capturing relationships between variables by mapping, interpreting, and modeling, (d) manipulating mathematical expressions and computing quantities, and (e) communicating these ideas in various forms.

The three quantitative literacy dimensions—problem-solving skills, content, and contexts—are shown in Tables 4, 5, and 6, respectively. Each of these aspects of the proposed quantitative literacy framework is aligned with definitions, frameworks, and assessments throughout the relevant quantitative literacy literature.

Mathematical Problem-Solving Skills

Employers need people "who can figure out how to solve a problem; that's more than just knowing how to plug numbers into a calculator in the right order" (Stumbo & Lusi, 2005, p. 1). When solving problems, examinees should be able to pose questions, organize information, draw diagrams, analyze situations, draw conclusions, and communicate and interpret results (Cohen, 1995). The five most common mathematical problem-solving skills identified among the existing quantitative literacy frameworks, definitions, and assessments are (a) interpretation, (b) strategic knowledge and reasoning, (c) modeling, (d) computation, and (e) communication. Table 4 defines each of these five skills briefly and describes foci for assessing each skill. For example, communication involves the presentation of mathematical concepts, data, procedures, and solutions in a variety of forms (e.g., written, graphic, or tabular format). Like other skills in the framework, communication is aligned with existing quantitative frameworks (i.e., AAC&U's Quantitative Literacy VALUE Rubric, AMATYC, MAA, and P21 Math) and assessments (i.e., CBAL, PIAAC Numeracy, and PISA Mathematics). These five problem-solving skills are not mutually exclusive but, instead, are dynamically interrelated. For instance, to appropriately interpret a graphical display, an examinee may be required to use strategic knowledge and reasoning skills to solve a problem.

Mathematical Content

Dwyer et al. (2003) noted that to solve quantitative reasoning problems, mathematical content knowledge is needed. After examining and synthesizing the quantitative literacy literature and existing assessments, we identified four content areas

Table 4 Description of Mathematical Problem-Solving Skill Areas for the Proposed Quantitative Literacy Framework

Problem-solving skill	Brief description	Focus of assessment
(a) Interpretation	The understanding and explanation of mathematical information, such as the ability to understand data, read graphs, draw conclusions, and recognize sources of error	*Understand mathematical terms and representational devices* Read and interpret basic mathematical notation, concepts, and terminology, such as percentage and average, as well as relationships between quantities expressed in terms of equations, formulas, or data representations, such as tables, graphs, and other diagrams
(b) Strategic knowledge and reasoning	The formulation and evaluation of mathematics problems using heuristics, and the ability to recognize relationships about mathematical concepts and situations	*Build and develop mathematical strategies* Construct and explore mathematical strategies and heuristics to solve problems using inductive and deductive reasoning *Develop and test conjectures* Formulate mathematical hypotheses and evaluate their consequences *Evaluate the validity of mathematical strategies* Evaluate the accuracy of solutions and detect any potential flaws or improbable results *Draw appropriate inferences and conclusions* Explain and justify mathematical results in different mathematical forms
(c) Modeling	The process of capturing relationships present in the environment or in mathematical forms, and expressing the model in one or more mathematical representations	*Translate information into mathematical forms* Convert informal contextual information into equations, graphs, diagrams, tables, or mathematical text *Map mathematical relationships* Use tools such as equations, inequalities, diagrams, two-way tables, graphs, flow-charts, and formulas to express quantitative relationships (e.g., linear relationships, triangle inequality) *Apply mathematical models* Apply mathematical models and relationships to real-world contexts
	The evaluation and revision of a model for accuracy and applicability	*Determine reasonableness of a mathematical model* Use estimation methods to check a solution; interpret the results and reflect on whether a solution makes sense *Revise mathematical models* Adjust mathematical models to make improvements if a model has not served its purpose
(d) Computation	The process of identifying and performing appropriate algebraic manipulations and arithmetic calculations to solve a problem	*Identify appropriate computational strategies* Correctly determine or select mathematical operations or computational methods for solving a problem *Accurately calculate* Perform arithmetic or algebraic operations accurately to solve a problem
(e) Communication	The presentation of higher-level concepts and ideas (e.g., mathematical arguments and models) as well as solutions to problems and more standard procedures; communication may take various mathematical forms and is customized to the appropriate target audience	*Present mathematical concepts, data, procedures, and solutions in a variety of forms* Communicate procedures and results in written, graphical, or tabular format using correct mathematical terminology and notation

Table 5 Mathematical Content Areas for the Proposed Quantitative Literacy Framework

Content area	Brief description	Focus of assessment
(a) Number and operations	Real numbers, order properties, and physical quantities	• Understand fundamental types of real numbers, including positive and negative numbers, integers, fractions and decimals, even and odd integers, prime numbers, rational and irrational numbers • Understand the order properties of real numbers and the number line • Understand physical quantities as real numbers with units, such as time money, weight, temperature, distance, area, and volume
	Arithmetic operations on real numbers	• Add, subtract, multiply, and divide real numbers, as well as exponentiate and take roots • Understand the properties of arithmetic operations (i.e., commutative, distributive) as well as the role the operations have in defining fractions, decimals, factors, multiples, and remainders • Understand relationships between arithmetic operations and the ordering of real numbers (e.g., the product of two negative numbers is a positive number)
	Estimation	• Use estimation to approximate answers • Use estimation to judge reasonableness of answers
	Proportional reasoning	• Compute and interpret percents and percent change • Compute and interpret rates, ratios, and proportions
(b) Algebra	Variables, algebraic expressions, and their use in representing quantities	• Use variables to represent varying quantities • Use arithmetic operations on variables to form algebraic expressions • Manipulate and simplify algebraic expressions
	Functions, their types and properties, and their use in solving problems	• Understand the concept of a function, including domain and range, use function notation, and evaluate functions • Know various types of elementary functions, including linear, quadratic, polynomial, and exponential • Understand properties of various types of functions • Represent and interpret functions graphically in a coordinate plane • Use functions to model varying quantities in order to solve problems
	Equations, inequalities, and their use in solving problems	• Understand equations and inequalities as conditions that must be satisfied by varying quantities • Solve problems using algebraic representations by setting up equations or inequalities involving functions or algebraic expressions • Graph equations and inequalities in a coordinate plane • Solve equations or inequalities algebraically, graphically, or by ad hoc methods, such as inspection or repeated substitution • Interpret solutions of equations or inequalities to solve problems

Table 5 Continued

Content area	Brief description	Focus of assessment
(c) Geometry and measurement	Geometric figures in one, two, and three dimensions	• Understand lines and angles in a plane, including parallel and perpendicular lines • Know two-dimensional and three-dimensional geometric figures, such as triangles, circles, polygons, rectangular solids, cylinders, and spheres • Understand transformations, congruence, and similarity of two-dimensional figures • Graph geometric figures in a coordinate plane
	Geometric figure measurements (e.g., area, distance, length, volume, angles) for solving a problem	• Calculate area and perimeter/circumference of a two-dimensional object • Calculate volume and surface area of a three-dimensional object • Measure angles of polygons • Use the Pythagorean theorem to calculate the side lengths of a triangle • Use measurement formulas (e.g., volume, area) to solve problems
	Units and systems of measurement	• Understand units of measurement (e.g., time, money, weight, temperature, distance, area, volume) and when to apply them • Make conversions within a system of measurement (e.g., inches to feet, meters to kilometers) • Convert from one system of measurement to another (e.g., US customary units to metric system, Fahrenheit to Celsius)
(d) Statistics and probability	Data interpretation and representation	• Read and interpret data in graphical or tabular form to solve problems • Determine appropriateness of a table or graph used to represent a set of data (e.g., line graphs vs. bar graphs) • Compare alternative displays of the same data set or displays across multiple data sets (e.g., bar graphs and pie graphs) for similarities and differences • Create a table to organize frequency data, proportional quantities, or the relationship between two variables • Represent the frequency distribution of data using a dotplot, histogram, boxplot, or stem-and-leaf plot • Plot proportional quantities using a pie or bar graph • Create line charts or scatterplots to represent the relationship between two variables
	Descriptive statistics	• Interpret and calculate measures of central tendency (e.g., mean, median, and mode) for a distribution of data • Interpret and calculate measures of dispersion or spread (e.g., standard deviation, range, interquartile range) for a distribution of data
	Basic probability	• Understand random sampling with and without replacement, and equal probability for all outcomes • Calculate the probability of a single event using fractions and proportions (e.g., the probability of selecting an ace in a deck of cards) • Calculate the probability of two (or more) independent events (e.g., probability of a coin coming up tails after two coin tosses) • Understand and calculate conditional probability (e.g., probability of selecting an ace on the second draw after selecting an ace on the first draw)

Table 6 Description of Real-World Contexts for the Proposed Quantitative Literacy Framework

Context	Examples
(a) Personal/everyday life	Handling money and budgets; shopping; time management; personal travel; playing games of chance; sports statistics; reading maps; measurement in cooking, home repairs, or personal hobbies; calculating a tip; completing an order form; understanding and evaluating personal health; balancing a checkbook
(b) Workplace	Managing schedules, budgets, and project resources; using spreadsheets; making and recording measurements; payroll/accounting; completing purchase orders; tracking expenditures; predicting costs; job-related decision making
(c) Society	Population changes; unemployment rates; voting systems; public transportation; government; public policies; demographics; advertising; national statistics; economics; quantitative information in the media; raising funds for an organization; interpretation of research studies; environmental trends or issues

important for assessing quantitative literacy in a higher education context: (a) number and operations, (b) algebra, (c) geometry and measurement, and (d) probability and statistics (see Table 5). As shown in Table 3, these content areas are aligned with many of the existing assessments measuring quantitative literacy skills. These four content areas are also aligned with the four strands of American Diploma Project benchmarks. These mathematics benchmarks were based on a review of day-to-day experiences of people in the workplace and college classrooms, such as courses in college algebra and calculus, introductory chemistry, or introductory microeconomics (Achieve, Inc., 2004).

In a higher education context, these different content areas should have varying weights throughout an assessment. In the K–12 Common Core State Standards for Mathematics, a clear shift is apparent in content emphasis through the grade levels, with basic number and operations being emphasized in the earlier K–5 grade levels while geometry, probability, and statistics are emphasized more throughout Grades 6–12. The understanding of basic probability and statistics has been identified as a requirement in today's informational and technological age (Stumbo & Lusi, 2005). Similarly, the understanding of numbers and numerical operations serves as the foundation for many entry-level positions in the workforce, with many jobs also requiring some working knowledge of algebra, geometry, data interpretation, and probability and statistics (Stumbo & Lusi, 2005). For instance, occupations such as machine operators and licensed nurses require knowledge in number sense and numerical operations, algebra, and geometry. Other jobs, such as actuaries or manufacturing technicians, require knowledge in data interpretation, probability, and statistics (Achieve, Inc., 2004).

Even though the proposed next-generation quantitative literacy assessment is intended for students in higher education, mathematical content learned in K–12 can provide the foundational knowledge for using real-world applications on a higher education assessment. Steen (2004) stated that many of the underpinnings of quantitative literacy are typically mathematical topics from middle school. The Common Core State Standards for Mathematics echoed the importance of middle school mathematics recognizing that "some of the highest priority content for college and career readiness comes from Grades 6–8" (NGA & CCSSO, 2010, p. 84). In a review of mathematical competencies focal to the middle grades, Graf (2009) summarized research that supports the notion that the connection between numbers and operations and algebra is difficult for middle school students to grasp and that this difficulty persists throughout high school and college. Therefore, K–12 mathematical content, especially concepts learned up until Grade 10, should serve as a foundation for developing a next-generation quantitative literacy assessment. The construct of quantitative literacy focuses on problem solving in applied and authentic contexts, requiring advanced reasoning skills rather than just relying on memorization skills. Using mathematical content up to Grade 10 encompasses the mathematical knowledge required of all students, regardless of college major or intended career path, not just those students in a mathematical-focused college major.

Real-World Contexts

Quantitative literacy involves the application of mathematical skills to solve real-world problems in varying contexts, situations, or settings (Dwyer et al., 2003; Rhodes, 2010). In the employer survey conducted by Hart Research Associates (2013), 78% of employers stated that more emphasis needs to be placed on a student's ability to apply knowledge to real-world settings. The American Diploma Project noted that contexts in postsecondary institutions and in the workplace are very different from the high school classroom (Achieve, Inc., 2004). The context of a mathematical problem impacts the strategy employed (e.g., operations and procedures) and level of knowledge required to comprehend and solve it. To apply the next-generation quantitative literacy assessment to the real world, the proposed framework uses three authentic contexts: (a) personal/everyday life, (b) workplace, and (c) society. The personal/everyday life context addresses topics such as handling money and budgets, shopping, calculating a tip, and balancing a checkbook. The workplace context addresses topics across a wide array of jobs, including tasks such as using spreadsheets, managing schedules and budgets, and predicting costs. Lastly, the society context addresses topics such as population and demographic changes, unemployment rates, and quantitative information found in the media. Table 6 lists these topics and additional topics for these three contexts. The contexts should have a broad scope to ensure that the assessment is fair for all students, regardless of college major.

The contexts align with contexts described in AAC&U's Quantitative Literacy VALUE Rubric, Lumina's DQP, P21 Math, Common Core State Standards for Mathematics, PIAAC Numeracy, and PISA Mathematics and have also been identified in the workforce literature. For example, in a survey by Casner-Lotto and Barrington (2006), 71.5% of the over 400 employer respondents identified "exercise in personal finance and responsibility, e.g., balancing a checkbook, budgeting skills, retirement planning" (p. 52) as one of the "most critical" (p. 51) emerging content areas for college graduates entering the workforce. Similarly, American Diploma Project identified that mathematical skills are needed across a variety of job categories to conduct tasks such as measuring and/or computing the concentration of a solution, measuring volume and weight, computing the dose of a medication, or computing cost estimates or credit requests (Achieve, Inc., 2004).

Assessment Structure: Item Formats and Task Types

To measure problem-solving skills and mathematics content within different contexts, a variety of item formats and task types should be considered. These item formats and task types should provide an authentic and interactive test-taking experience for examinees. This section provides recommendations for both item formats and task types when developing a next-generation quantitative literacy assessment. It is recommended that all item formats employ the use of automated scoring because of its efficiency in providing instantaneous feedback to examinees (Zhang, 2013).

Item Formats

Most existing assessments measuring quantitative literacy skills in higher education use two item formats: selected-response and open-ended items. Examples of selected-response items include multiple-choice items or drop-down menu items. Similarly, examples of open-ended items include numeric entry or short answer items. Among the existing assessments, many assess students using selected-response items, specifically, multiple-choice items, because they are easy to score and typically yield higher reliability estimates. Typically, with selected-response items, more items can be administered in a shorter period of time, also allowing for more content coverage (R. E. Bennett, Morley, & Quardt, 2000). Existing assessments have also used some open-ended items such as numeric entry. Open-ended formats can eliminate random guessing and reduce measurement error, eliminate corrective feedback, and remove an examinee's ability to work backward from answer choices (Bridgeman, 1993). Table 7 shows a list of recommended item formats to be used in a next-generation quantitative literacy assessment that allows for automated scoring.

Selected-response items include both single- or multiple-selection items and computer-facilitated item formats such as clicking/selecting objects, drop-down menus, and drag and drop. For instance, a clicking/selecting objects item, similar to those found on PIAAC Numeracy, could be used to assess problem-solving skills such as interpretation or modeling. Additionally, a drop-down menu could be used to communicate information in graphical form. Drag and drop items could be used to correctly order a series of numbers. Open-ended items include numeric and expression entry, create/edit a graph/table, and short constructed response. Unlike numeric entry, expression entry (i.e., mathematical expression or a generating example item) allows an examinee to create a mathematical model. To avoid input errors, an expression item should have specific buttons that allow the examinee to develop the expression (R. E. Bennett et al., 2000). The create/edit a graph/table is similar to the graphical modeling item discussed in R. E. Bennett et al., where an examinee plots points and uses a tool to connect all of plotted points. These recommended selected-response and open-ended items can be technologically innovative and promote higher-level thinking.

Table 7 Next-Generation Quantitative Literacy Assessment Item Formats

Item format category[a]	Format response type	Description
Selected response	Single- and multiple-selection multiple choice	A question/stem with multiple answer choices of which one or more could be a correct response
	Drop-down menu	A variation of a traditional multiple-choice item, where one answer choice is selected via a drop-down menu; can use when customizing figures and tables
	Clicking/selecting objects	Select/click/highlight one or more parts of a table, figure, or text to answer the question
	Drag and drop	An examinee selects objects and places them in a specific location or order
Open ended	Numeric entry	A numerical answer must be entered rather than selected from a list of answer choices
	Expression entry	A mathematical expression must be entered using buttons that represent numbers and operators, variables, constants, and sub- and super-scripts
	Create/edit a graph/table	An examinee uses information in the question/stem to develop/edit a graphical display or table
	Short constructed response	Examinees must respond in their own words to a prompt based on text, graph, or other stimuli

[a] Items may be discrete or grouped together with a common figure or table.

It is important to understand that item format can impact both item difficulty and the reliability of the test. For example, Bridgeman (1993) found that when using the same underlying test questions in different formats (i.e., multiple-choice vs. numeric or simple formula entry), some items that were relatively easy in a multiple-choice format were quite difficult for examinees in a numeric entry format. The author also noted that the reliability of an assessment with either all multiple-choice or all numeric entry items can vary when the assessment has a fixed number of items or fixed testing time. For instance, with a fixed number of test items, an assessment with all numeric entry items may be more reliable because random guessing is removed (Bridgeman, 1993). However, it is important to note that within a fixed testing time, an assessment with all multiple-choice items is more likely to yield higher reliability estimates when compared to an assessment with all numeric entry items because the numeric entry item assessment would have fewer test items.

Task Types

Task types are used to address the foci of the assessment and can be used to enhance the previously described item formats. Table 8 provides a list of suggested task types for a next-generation quantitative literacy assessment. Similar tasks can be found on assessments such as CBAL Mathematics, GMAT Integrative Reasoning and GMAT Quantitative, CLA+ SQR, EPP Mathematics, PISA Mathematics, and GRE Quantitative Reasoning. Task types, in combination with various item formats, allow for more authenticity and examinee engagement. For instance, a task with a “clicking objects” response format (see Table 7) could assess whether an examinee can recognize an inconsistency among formulas in a spreadsheet (e.g., if a cell containing a formula for the sum of column totals and a cell containing a formula for the sum row totals from the same table do not match). The student could use the computer mouse to click in the cell or cells within the spreadsheet where the errors occur. This combination of item format and task type can recreate a real-world situation for examinees.

A Sample Item

Some of the considerations for a next-generation assessment framework can be clarified by a sample item that fits into each of three dimensions: problem-solving skill, content area, and context. Figure 1 shows a sample item written in the context of personal/everyday life (determining the data plan for a cell phone) and requires the examinee to use modeling, one of the problem-solving skills identified in Table 4. Part A of the problem requires the examinee to translate information into mathematical forms, and Part B requires the examinee to apply that mathematical model. In terms of content, this problem uses algebraic expressions and functions (see Table 5). The item response formats (Table 7) for this problem

Table 8 Next-Generation Quantitative Literacy Assessment Task Types

Task type	Description
Data accuracy	Statements are provided in reference to a question/stem. Examinee must determine whether the statements are accurate with respect to the stem
Data sufficiency	Statements are provided in reference to a question/stem. Examinee must determine whether the statements provide sufficient information to solve a problem or answer a question
Draw conclusions	Examinee reads data and draws conclusions or computes quantities
Evidence	Examinee provides or identifies types of information that must be sought in order to evaluate an argument, statement, or claim
Quantitative comparison	Examinee compares two presented quantities (less than, equal to, or greater than) or determines that there is not enough information to make the comparison
Recognize inconsistency	Examinee recognizes inconsistencies or errors in mathematical information (e.g., statements, equations, figures, or tables)
Representational equivalence	Examinee makes comparisons between two graphs, two tables, or a graph and a table and determines whether they are equivalent

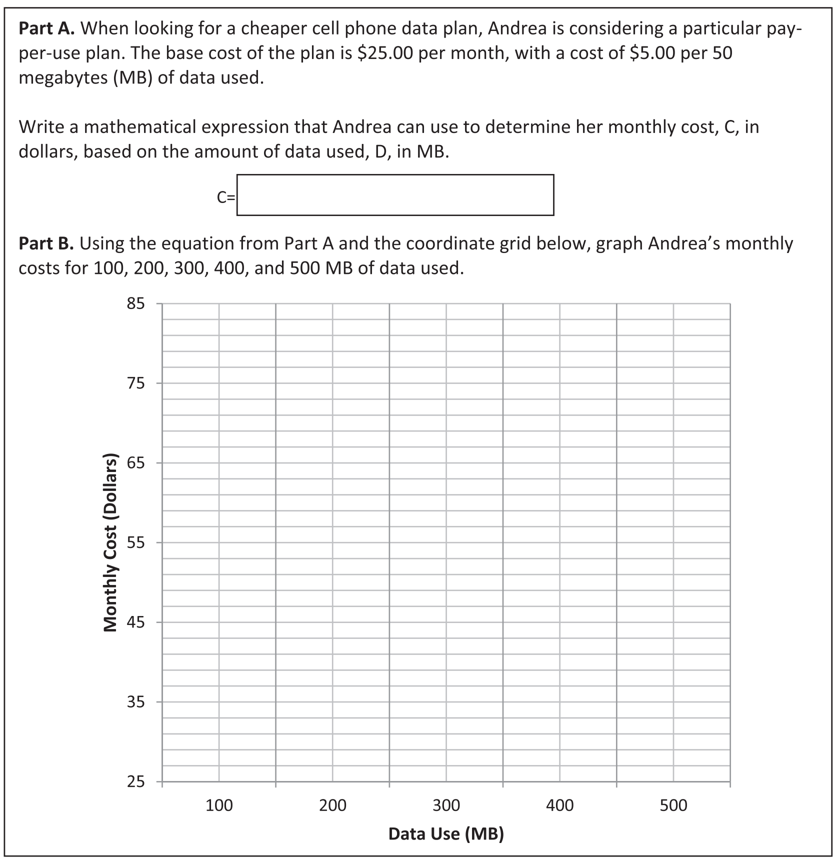

Figure 1 Sample expression entry item (Part A) and create a graph item (Part B).

include expression entry (Part A) and create a graph (Part B), and the item task type (Table 8) is draw conclusions. Other ways to represent this item without changing the construct, especially for Part B, include the use of drag and drop or clicking item formats. For instance, the examinee could drag a function line that is already created and place it on the graph or click on the function line. Both Parts A and B could also use multiple-choice items, but the use of alternative item formats makes the assessment more authentic. That said, because the item can use a variety of interaction methods, it is potentially more accessible for all students.

Potential Sources of Construct-Irrelevant Variance

When developing a next-generation quantitative literacy assessment in higher education, it is important to consider the accessibility of the assessment to all students, including students with limited English proficiency or students with disabilities. Developing an accessible assessment means careful consideration of multiple modes of delivery, as well as accessible methods for accessing test questions and entering responses. Technology-enhanced item formats may open doors for some students, while creating barriers for others, especially for students not used to completing quantitative problems on a computer. The choice of item formats (i.e., technology-enhanced or traditional) should depend on a number of factors such as construct representation, target population, scoring accuracy, scoring efficiency, and testing time. New item formats need to have clear directions on how to approach those items. For example, a clicking/selecting objects item could have the following instructions: Click two bars on the bar graph to answer the question below. The instructions would also need to indicate how the bar might change to show that it has been selected, and would need to indicate how to unselect the bar. A clear instruction is needed to ensure that examinees know how to appropriately approach the problem.

Another potential source of construct-irrelevant variance is the cognitive reading load on test questions. When assessing quantitative literacy, there is a certain level of assumed reading ability required to understand and answer test questions (Dwyer et al., 2003), especially as the items focus more on reasoning and interpretation skills. Research in K–12 educational assessment has investigated the impact of language factors on mathematics assessments. Abedi and Lord (2001) found that Grade 8 English language learners (ELLs) scored significantly lower than proficient English speakers on mathematical word problems and that ELLs performed better when taking the assessment with linguistically simplified test items. The authors noted that the complexity of text in mathematics problems was a larger issue for inexperienced and nonexpert problem solvers. Shaftel, Belton-Kocher, Glasnapp, and Poggio (2006) also evaluated the impact of linguistic features on mathematics items, finding a significant impact on student performance in Grades 4, 7, and 10. Interestingly, however, the effect size decreased as grade level increased, and Grade 10 item difficulty was most influenced by difficult mathematics vocabulary rather than other linguistic characteristics (Shaftel et al., 2006). It is likely that at the college level, mathematics vocabulary will also play a significant role in the difficulty of mathematical word problems. Therefore, the level of English proficiency required to answer a question should be considered in the development of test items.

Potential Advantages of the Proposed Framework and Assessment Considerations

The proposed framework and assessment considerations offer some distinct advantages over existing frameworks and assessments. A common theme throughout the literature on quantitative literacy and related constructs (e.g., quantitative reasoning) is that a quantitatively literate person needs to have the ability to interpret and manipulate mathematical information in many different contexts such as personal, work, and civic lives (Gillman, 2006). Many of the existing frameworks and assessments stress the importance of application to real-world problems but do not provide specific contexts (e.g., AAC&U's VALUE Rubric, Lumina's DQP, USDOL).The proposed framework has the advantage of defining three contexts for writing test items. The proposed framework also clearly defines both problem-solving skills and content areas, as both are important in developing a next-generation quantitative literacy assessment. Many of the existing frameworks and assessments (e.g., AAC&U's Quantitative Literacy VALUE Rubric) clearly define problem-solving skills but do not define mathematical content. Other frameworks and assessments (e.g., CAAP Mathematics) clearly define mathematical content and not the problem-solving skills.

Another advantage of the proposed framework is the alignment with existing frameworks, definitions, and assessments. The proposed framework carefully synthesizes problem-solving skills, content, and contexts identified by national and international organizations, workforce initiatives, higher education institutions and researchers, and K–12 theorists and practitioners. This synthesized approach is unique to the proposed framework. Additional advantages of the proposed framework and assessment considerations are the use of computer-facilitated item formats and the consideration of accessibility when developing the assessment.

Conclusion

Quantitative literacy has been identified as an important learning outcome to both higher education institutions and employers, making it critical to evaluate whether college students receive sufficient training on quantitative skills throughout their postsecondary education. A review of frameworks by various stakeholders and a review of existing assessments showed that many common themes exist when defining quantitative literacy or other related terms such as quantitative reasoning and numeracy. These common themes include overlap among mathematical skills, content areas, and contexts. This research also stressed that quantitative literacy involves more application and problem solving and less strict memorization of mathematical content. The existing research and assessments informed both the proposed framework and proposed assessment that involve applications and problem-solving skills across mathematical content areas. We hope this proposed framework provides clear guidance to stakeholders and assessment developers about the important aspects of the quantitative literacy construct. Using this framework to guide the development of a next-generation quantitative literacy assessment requires coordination between content experts, assessment developers, measurement experts, and potential score users at higher education institutions. This coordinated assessment development process requires both careful test development and psychometric evaluation to create a quality learning outcomes assessment for measuring quantitative literacy at the college level.

Acknowledgments

The authors would like to thank Daryl Ezzo for sharing his knowledge of mathematics assessment development and for assisting with the proposed framework and assessment considerations. The authors would also like to thank James Wyman Brantley for providing feedback and edits on this report during its development.

References

Abedi, J., & Lord, C. (2001). The language factor in mathematics tests. *Applied Measurement in Education*, *14*(3), 219–234.

Abel, J. R., & Deitz, R. (2013). *Do big cities help college graduates find better jobs?* Retrieved from http://libertystreeteconomics.newyorkfed.org/2013/05/do-big-cities-help-college-graduates-find-better-jobs.html

Achieve, Inc. (2004). *Ready or not: Creating a high school diploma that counts*. Washington, DC: Author.

Adelman, C., Ewell, P., Gaston, P., & Schneider, C. G. (2011). *The degree qualifications profile*. Indianapolis, IN: Lumina Foundation.

Adelman, C., Ewell, P., Gaston, P., & Schneider, C. G. (2014). *The degree qualifications profile 2.0: Defining U.S. degrees through demonstration and documentation of college learning*. Indianapolis, IN: Lumina Foundation.

Admissions Testing Service. (2013). *About STEP*. Retrieved from http://www.admissionstestingservice.org/our-services/subject-specific/step/about-step/

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (1999). *The standards for educational and psychological testing*. Washington, DC: American Educational Research Association.

Association of American Colleges and Universities. (2011). *The LEAP vision for learning: Outcomes, practices, impact, and employers' view*. Washington, DC: Author.

Baldi, S. (2009). *Technical report and data file user's manual for the 2003 national assessment of adult literacy* (NCES 2009-476). Washington, DC: US Department of Education, National Center for Education Statistics.

Bennett, J., & Briggs, W. L. (2008). *Using and understanding mathematics: A quantitative reasoning approach* (4th ed.). Reading, MA: Addison.

Bennett, R. E., Morley, M., & Quardt, D. (2000). Three response types for broadening the conception of mathematical problem solving in computerized tests. *Applied Psychological Measurement*, *24*(4), 294–309.

Bridgeman, B. (1993). *A comparison of open-ended and multiple-choice question formats for the quantitative section of the graduate record examinations general test* (GRE Board Professional Report No. 88-13P). Princeton, NJ: Educational Testing Service.

Bridgeman, B., Harvey, A., & Braswell, J. (1995). Effects of calculator use on scores on a test of mathematical reasoning. *Journal of Educational Measurement*, *32*(4), 323–340.

CAAP Program Management. (2012). *ACT CAAP technical handbook 2011–2012*. Iowa City, IA: Author.

Capital Workforce Partners. (2014). *The career competencies system*. Retrieved from http://capitalworkforce.org/youth-resources/career-competencies-system/

Carleton College. (2013). *Carleton's quantitative inquiry, reasoning, and knowledge (QuIRK) initiative*. Retrieved from http://serc.carleton.edu/quirk/index.html

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 U.S. workforce*. New York, NY: Conference Board, Partnership for 21st Century Skills, Corporate Voices for Working Families, Society for Human Resource Management.

Cohen, D. (Ed.) (1995). *Crossroads in mathematics: Standards for introductory college mathematics before calculus*. Memphis, TN: American Mathematical Association of Two-Year Colleges.

College Board. (2012). *Test information guide: College-Level Examination Program® 2012–13 college mathematics*. New York, NY: Author.

Dwyer, C. A., Gallagher, A., Levin, J., & Morley, M. E. (2003). *What is quantitative reasoning? Defining the construct for assessment purposes* (Research Report No. RR-03-30). Princeton, NJ: Educational Testing Service.

Educational Testing Service. (2010). *ETS proficiency profile user's guide*. Princeton, NJ: Author.

Educational Testing Service. (2013a). *GRE guide to the use of scores 2013–2014*. Princeton, NJ: Author.

Educational Testing Service. (2013b). *GRE test content and structure*. Princeton, NJ: Author.

Educational Testing Service. (2013c). *Introduction to the quantitative reasoning measure*. Retrieved from http://www.ets.org/gre/revised_general/prepare/quantitative_reasoning

Educational Testing Service. (2014a). *The CBAL™ initiative: Innovation in K–12 assessment*. Retrieved from http://www.ets.org/research/topics/cbal/initiative/

Educational Testing Service. (2014b). *GRE: Using a calculator*. Retrieved from https://www.ets.org/gre/revised_general/prepare/quantitative_reasoning/calculator

Ellington, A. J. (2003). A meta-analysis of the effects of calculators on students' achievement and attitude levels in precollege mathematics classes. *Journal for Research in Mathematics Education*, *34*(5), 433–463.

Estry, D. W., & Ferrini-Mundy, J. (2005). *Quantitative literacy task force final report and recommendations*. Retrieved from http://acadgov.msu.edu/documents/QLFinalReport-1.pdf

Ewell, P. T. (2001). Numeracy, mathematics, and general education. In L. A. Steen (Ed.), *Mathematics and democracy: The case for quantitative literacy* (pp. 37–48). Princeton, NJ: Woodrow Wilson National Fellowship Foundation.

Gillman, R. (Ed.) (2006). *Current practices in quantitative literacy*. Washington, DC: Mathematical Association of America.

Goodman, M., Finnegan, R., Mohadjer, L., Krenzke, T., Hogan, J., Owen, E., & Provasnik, S. (2013). *Literacy, numeracy, and problem-solving technology-rich environments among U.S. adults: Results from the Program for the International Assessment of Adult Competencies 2012* (NCES 2014-008). Washington, DC: U.S. Department of Education, National Center for Educational Statistics.

Graduate Management Admission Council. (2013a). *Quantitative section*. Retrieved from http://www.mba.com/us/the-gmat-exam/gmat-exam-format-timing/quantitative

Graduate Management Admission Council. (2013b). *Validity, reliability, & fairness*. Retrieved from http://www.gmac.com/gmat/the-gmat-advantage/validity-reliability-fairness.aspx

Graf, E. A. (2009). *Defining mathematics competency in the service of cognitively based assessment for grades 3 through 8* (Research Report No. RR-09-42). Princeton, NJ: Educational Testing Service.

Hart Research Associates. (2013). *It takes more than a major: Employer priorities for college learning and student success*. Washington, DC: Author.

Hollins University. (2013). *Course goals and criteria document*. Retrieved from https://www.hollins.edu/registrar/documents/ESP_goalscriteria.pdf

Insight Assessment. (2013a). *Quant Q: Measure quantitative reasoning*. San Jose: California Academic Press.

Insight Assessment. (2013b). *Test of everyday reasoning-numeracy (TER-N)*. San Jose: California Academic Press.

Institute of Educational Statistics. (n.d.). *National Assessment of Adult Literacy (NAAL)*. Retrieved from http://nces.ed.gov/naal/si_questions.asp

Kirsch, I. S., Jungeblut, A., Jenkins, L., & Kolstad, A. (2002). *Adult literacy in America: A first look at the findings of the national adult literacy survey* (NCES 1993-275). Washington, DC: U.S. Department of Education.

Klein, S., Liu, O. L., Sconing, J., Bolus, R., Bridgeman, B., Kugelmass, H., … Steedle, J. (2009). *Test Validity Study (TVS) report*. Retrieved from https://cp-files.s3.amazonaws.com/26/TVSReport_Final.pdf

Kuncel, N. R., Hezlett, S. A., & Ones, D. S. (2001). A comprehensive meta-analysis of the predictive validity of the graduate record examinations: Implications for graduate student selection and performance. *Psychological Bulletin*, *127*(1), 162–181.

Lakin, J. M., Elliott, D. C., & Liu, O. L. (2012). Investigating ESL students' performance on outcomes assessments in higher education. *Educational and Psychological Measurement*, *72*(5), 734–753.

Langkamp, G., & Hull, J. (2007). *Quantitative reasoning and the environment*. Upper Saddle River, NJ: Pearson Education.

Liu, O. L., Bridgeman, B., & Adler, R. M. (2012). Measuring learning outcomes in higher education: Motivation matters. *Educational Researcher*, *41*(9), 352–362.

Liu, O. L., & Roohr, K. C. (2013). Investigating 10-year trends of learning outcomes at community colleges (Research Report No. RR-13-34). Princeton, NJ: Educational Testing Service.

Lutzer, D. J., Rodi, S. B., Kirkman, E. E., & Maxwell, J. W. (2007). *Statistical abstract of undergraduate programs in the mathematical sciences in the United States: Fall 2005 CBMS survey*. Providence, RI: American Mathematical Society.

Marr, D. (1995). *Validity of the academic profile*. Princeton, NJ: Educational Testing Service.

Mayes, R. L., Peterson, F., & Bonilla, R. (2013). Quantitative reasoning for learning progressions for environmental science: Developing a framework. *Numeracy*, *6*(1), Article 4. doi:10.5038/1936-4660.6.1.4

McKinsey & Company. (2013). *Voice of the graduate*. Philadelphia, PA: Author.

Merriam-Webster. (2014). Literacy. Retrieved from http://www.merriam-webster.com/dictionary/literacy

Mount St. Mary's College. (2013). *Definition of quantitative literacy*. Retrieved from http://www.msmc.la.edu/academics/quantitative-literacy-program/definition-of-quantitative-literacy.asp

National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010). *Common core state standards for mathematics*. Washington, DC: Author.

National Numeracy Network. (2013). *What is numeracy/QL/QR?* Retrieved from http://serc.carleton.edu/nnn/resources/index.html

National Survey of Student Engagement. (2013a). *A fresh look at student engagement: Annual results 2013*. Bloomington: Indiana University Center for Postsecondary Research.

National Survey of Student Engagement. (2013b). *NSSE 2013 U.S. grand frequencies: Frequency distributions by Carnegie classification—First-year students*. Retrieved from http://nsse.iub.edu/2013_institutional_report/pdf/Frequencies/Freq%20-%20FY%20by%20Carn.pdf

National Survey of Student Engagement. (2013c). *NSSE 2013 U.S. grand frequencies: Frequency distributions by Carnegie classification—Seniors*. Retrieved from http://nsse.iub.edu/2013_institutional_report/pdf/Frequencies/Freq%20-%20SR%20by%20Carn.pdf

Organisation for Economic Co-Operation and Development. (2012a). *Literacy, numeracy and problem solving in technology-rich environments: Framework for the OECD survey of adult skills*. Retrieved from http://www.oecd.org/site/piaac/PIAAC%20Framework%202012--%20Revised%2028oct2013_ebook.pdf

Organisation for Economic Co-Operation and Development. (2012b). *The survey of adult skills: Reader's companion*. Paris, France: Author.

Organisation for Economic Co-Operation and Development. (2013). *PISA 2012 assessment and analytical framework: Mathematics, reading, science, problem solving, and financial literacy*. Paris France: Author.

Organisation for Economic Co-Operation and Development and Statistics Canada. (2000). *Literacy in the information age. Final report of the international adult literacy survey*. Paris, France: Author.

Partnership for Assessment of Readiness for College and Careers. (2014). *The PARCC assessment*. Retrieved from http://www.parcconline.org/parcc-assessment

PIAAC Numeracy Expert Group. (2009). *PIAAC numeracy*: *A conceptual framework* (OECD Education Working Paper No. 35). Paris, France: OECD.

Rhodes, T. L. (Ed.) (2010). *Assessing outcomes and improving achievement: Tips and tools for using rubrics*. Washington, DC: Association of American Colleges and Universities.

Saltrick, S., Hadad, R., Pearson, M., Fadel, C., Regan, B., & Winn, J. (2011). *21st century skills map*. Washington, DC: Partnership for 21st Century Skills Retrieved from http://www.p21.org/storage/documents/P21_Math_Map.pdf

Samford University. (2009). *Quantitative literacy*. Retrieved from http://www.samford.edu/ctls/archives.aspx?id=2147484121

Shaftel, J., Belton-Kocher, E., Glasnapp, D., & Poggio, J. (2006). The impact of language characteristics in mathematics test items on the performance of English language learners and students with disabilities. *Educational Assessment*, *11*(2), 105–126.

Sinharay, S., Puhan, G., & Haberman, S. J. (2011). An NCME instructional module on subscores. *Educational Measurement: Issues and Practice*, *30*(3), 29–40.

Smarter Balanced Assessment Consortium. (n.d.). *Smarter balanced assessments*. Retrieved from http://www.smarterbalanced.org/smarter-balanced-assessments/

Sons, L. (Ed.) (1996). *Quantitative reasoning for college graduates: A complement to the standards*. Washington, DC: Mathematical Association of America.

Steele, B., & Kilic-Bahi, S. (2008). Quantitative literacy across the curriculum: A case study. *Numeracy*, *1*(2), 1–15.

Steen, L. A. (2001). The case for quantitative literacy. In L. A. Steen (Ed.), *Mathematics and democracy* (pp. 1–22). Princeton, NJ: Woodrow Wilson National Fellowship Foundation.

Steen, L. A. (2004). *Achieving quantitative literacy: An urgent challenge for higher education*. Washington, DC: Mathematical Association of America.

Stumbo, C., & Lusi, S. F. (2005). *Why isn't the mathematics we learned good enough for today's students? The imperative of mathematical literacy*. Washington, DC: Council of Chief State School Officers and Texas Instruments.

Sundre, D. L. (2008). *The Scientific Reasoning Test, Version 9 (SR-9) test manual*. Harrisonburg, VA: Center for Assessment and Research Studies.

Tamassia, C., Lennon, M. L., & Yamamoto, K. (2013). Scoring and reliability studies. In I. Kirsch and W. Thorn (Eds.), *Technical report of the survey of adult skills (PIAAC)*. Paris, France: OECD.

Traub, R. E., & Rowley, G. L. (1991). An NCME instructional module on understanding reliability. *Educational Measurement: Issues and Practice*, *10*(1), 37–45.

U.S. Department of Labor. (2013). *Competency model clearinghouse: Mathematics*. Retrieved from http://www.careeronestop.org/competencymodel/blockModel.aspx?tier_id=2&block_id=9

University of Kentucky. (2012). *UK core quantitative reasoning rubric*. Retrieved from http://www.uky.edu/ukcore/sites/www.uky.edu.ukcore/files/QR%20Rubric.pdf

University of Warwick. (2013). *STEP papers and the advanced extension award in mathematics*. Retrieved from http://www2.warwick.ac.uk/fac/sci/maths/admissions/ug/aeastep/

Vacher, H. L. (2014). Looking at the multiple meanings of numeracy, quantitative literacy, and quantitative reasoning. *Numeracy*, *7*(2), 1–14.

Voluntary System of Accountability. (2013). *Voluntary system of accountability program*. Retrieved from http://www.voluntarysystem.org/slo_evaluation

Ward, R. M., Schneider, M. C., & Kiper, J. D. (2011). Development of an assessment of quantitative literacy for Miami University. *Numeracy*, *4*(2), Article 4. doi:10.5038/1936-4660.4.2.4

Wellesley College. (n.d). *Study packet for the quantitative reasoning assessment*. Wellesley, MA: Author.

Zahner, D. (2013). *Reliability and validity—CLA+*. New York, NY: Council for Aid to Education.

Zhang, M. (2013). Contrasting automated and human scoring of essays. *R&D Connections*, *21*, 1–11.

Suggested citation:

Roohr, K. C., Graf, E. A., & Liu, O. L. (2014). *Assessing quantitative literacy in higher education: An overview of existing research and assessments with recommendations for next-generation assessment* (ETS Research Report No. RR-14-22). Princeton, NJ: Educational Testing Service. doi:10.1002/ets2.12024

Action Editor: Donald Powers

Reviewers: James Brantley and James Fife