Adult numeracy and its assessment in the ALL survey: A conceptual framework and pilot results

Aug 2003

Written by members of the ALL Numeracy Team
(in alphabetical order):

Iddo Gal, University of Haifa, Israel
Mieke van Groenestijn, Utrecht University of Professional Education, Netherlands
Myrna Manly (ret.), El Camino College, California, USA
Mary Jane Schmitt, TERC, Cambridge, Massachusetts, USA
Dave Tout, Language Australia, Melbourne, Australia.

Team Liaison and consultant for Statistics Canada: Yvan Clermont, Statistics Canada.

Source: Statistics Canada, Ottawa, Canada.

Available online: www.ets.org/all

Overview

About the ALL Survey

The Adult Literacy and Lifeskills (ALL) survey (formerly the International Life Skills Survey, ILSS), is a follow-up to the International Adult Literacy Survey (IALS), the world's first large-scale comparative assessment of adult literacy. The ALL survey has been a joint development by Statistics Canada and by the United States' National Center for Education Statistics (NCES), in cooperation with the Organization for Economic Cooperation and Development (OECD) and the United Nations Educational, Scientific, and Cultural Organization (UNESCO (OREALC)) since 1997.

Using household survey methods, the ALL project assesses performance of adults aged 16-65 in the domains of *Prose* and *Document Literacy*, *Numeracy*, and *Problem Solving*, and also collects information about experiences with Information and Communication Technology. Many other variables and correlates of interest, such as demographic details, employment status, and participation in learning activities, are recorded via a Background Questionnaire (BQ).

The key goals of the project are to:

- Profile and compare skill distributions across participating countries
- Explore covariates of observed skills, e.g., social and economic factors,
- Understand the relationship of Numeracy and Problem Solving to Prose and Document Literacy.

Following several years of preparation and an extensive Pilot survey in 2002, the first round of ALL's main assessment, involving 6 countries, began in 2003. Plans include a full comparative report from the first round, to be followed by other publications and by assessment rounds in additional countries.

Numeracy in the ALL survey

Numeracy is included as a domain in the ALL Survey as one of the critical factors in determining the capability of a population to adapt to and effectively function in an increasingly information-laden society or to perform well at work (European Commission, 1996). Schools are placing more emphasis on the links between the knowledge and skills gained in the mathematics classroom and students' ability to handle real-life situations that require activation of mathematical knowledge and skills. Given the increasing need for adults to continuously adapt to changing citizenship, workplace, and everyday life demands, it is vital that nations have information about their workers' and citizens' numeracy in order to evaluate the human capital available for advancement, to plan effective school-based and lifelong learning opportunities, and to better understand the factors that affect citizens' ability to advance their well-being.

The conception of numeracy developed for ALL is built upon recent research and work done in several countries on functional demands of different life contexts, on the nature of adults' mathematical and statistical knowledge and skills, and on how such skills are applied or used in different circumstances. In light of the general intention of the ALL survey to provide information about a diverse set of lifeskills, this framework defines numeracy as follows:

Numeracy is the knowledge and skills required to effectively manage and respond to the mathematical demands of diverse situations.

However, since an assessment can only examine observed behavior, not internal processes or capacities, this framework uses a more detailed definition of "numerate behavior" as a means to guide the development of items for the survey.

Numerate behavior is observed when people manage a situation or solve a problem in a real context; it involves responding to information about mathematical ideas that may be represented in a range of ways; it requires the activation of a range of enabling knowledge, factors, and processes.

Both definitions above use broad concepts and terms whose specific meanings and underlying components are explained in more detail later in this document. The definitions imply that numeracy can be viewed as a functional competency that is somewhat different from the traditional notion of "knowing school mathematics" in that it relates to the capacity to act and bring one's knowledge (mathematical and other) to bear on tasks *in context.*

Numeracy is assessed in the ALL survey by presenting to respondents short tasks with mathematical content that are embedded in real-life contexts. As illustrated by the sample items in Appendix 1, tasks require that respondents activate a range of knowledge and skills, and respond to different situations by computing, estimating, understanding notions of shape, length, volume, monetary units, measuring, understanding some statistical ideas, or interpreting simple formulas. Respondents are encouraged to use the tools provided, a ruler and a four-function calculator, whenever they wish. Numeracy and its assessment is thus broader in scope than Quantitative Literacy as defined in the IALS, which refers to a person's ability to apply *arithmetic* operations to numbers embedded in print materials.

About this report

This report describes key stages in the development of the Numeracy assessment scale for the ALL survey that took place between 1998 and early 2003, in three parts:

Part A presents a conceptual discussion of the numeracy construct and its facets, examines approaches to assessment of mathematical skills, and reviews issues that influenced the item development process. This section includes most of the text that appeared in the original Numeracy framework published in 1999 on the ALL (then ILSS) website. Editorial changes were made to streamline the presentation and to respond to comments from external reviewers, but the material was kept mostly intact since it served as the conceptual foundation from which item development progressed starting in 1998.

Part B describes the development of the item pool and scoring guidelines, and the feasibility studies that led to the selection of 80 items for a Pilot study that took place in 2002.

Part C outlines the design of the ALL Pilot study and presents key results, on the basis of which 40 Numeracy items were selected for the main ALL assessment in 2003. Two *appendices* contain sample Numeracy items that have been released to the public (Appendix 1) and details of a scheme of complexity factors that was used to inform the evaluation of difficulty levels of items during the item development process (Appendix 2).

Table of Contents

Acknowledgments

We thank the numerous international experts, study managers, and translators who have commented on earlier drafts of this framework or on sample items, or who have made suggestions regarding assessment tasks, items, and procedures.

Part A: Conceptual foundations

This part of the report presents a framework for the assessment of numeracy skills in the Adult Literacy and Lifeskills Survey (ALL). It begins by examining different perspectives that informed the conceptualization of numeracy developed for the ALL survey, and reviews some factors known to affect adults' numerate behavior, such as literacy, beliefs and attitudes, and prior practices, some of which are studied by specific questions in the ALL Background Questionnaire (BQ). Next, a definition of numeracy is presented, followed by an elaboration on the facets of numerate behavior that guided the development of items for the ALL Numeracy scale. Finally, several factors that are thought to influence the complexity or difficulty of numeracy tasks are presented; these factors are important both for development of items as well as for interpretation of the survey results. The actual stages of item development and pilot testing are described in parts B and C of the report.

1. Adult Numeracy: Influences and Perspectives

The construct "numeracy" does not have a universally accepted definition, nor agreement about how it differs from "mathematics." For some educators and officials, the term numeracy implies a set of simple skills involving the ability to carry out basic computations or arithmetical operations. In contrast, adult numeracy as viewed here is a broad construct that involves a range of knowledge, skills, and supporting processes. This section presents a review of influences or trends that have led to the conceptualization of numeracy employed in the ALL survey, and organizes them under five headings: *Workplace perspectives, Broader life purposes, Educational perspectives, Research perspectives, and Assessment schemes*. (These five headings or areas are interconnected and not mutually exclusive, as developments in one area often influence thinking in other areas).

1.1 Workplace perspectives

Over the last two decades, many countries have launched large-scale efforts aimed to define "core skills" or "key competencies" that workers should have, in response to the need to maintain economic competitiveness and improve employability of adults and school graduates. As workplaces are becoming more concerned with involving all workers in improving workplace efficiency and quality processes, the importance of numeracy skills is growing, and they have been shown to be a key factor in workplace success (Jones, 1995; Murnane, Willett & Levy, 1995). Basic computational knowledge has always been considered as part of the fundamental skills that adults need to possess, but the recent skills frameworks describe in specific terms the need for workers to possess a much broader range of mathematical skills. Examples exist in many countries, such as the United Kingdom, the United States, Australia, the Netherlands, and other OECD countries. The following descriptions are indicative of the nature of such efforts.

Outcomes from skills projects conducted in the United States illustrate workplace perspectives regarding mathematical skills needed by workers. Following earlier research by a task force of the American Society of Training and Development (Carnevale, Gainer, & Meltzer, 1990), the Secretary of Labor’s Commission on Achieving Necessary Skills (SCANS) (Packer, 1997) has differentiated between mastery of basic arithmetical skills and much broader and flexible understanding of principles and underlying ideas, subsumed under the notion of mathematical skills:

> *SCANS arithmetical skills:* Performs basic computations; uses basic numerical concepts such as whole numbers and percentages in practical situations; makes reasonable estimates and arithmetic results without a calculator; and uses tables, graphs, diagrams and charts to obtain or convey quantitative information.
>
> *SCANS mathematical skills:* Approaches practical problems by choosing appropriately from a variety of mathematical techniques; uses quantitative data to construct logical explanations for real world situations; expresses mathematical ideas and concepts orally and in writing; and understands the role of chance in the occurrence and prediction of events. (SCANS, 1991, p. 83)

Based on a later survey of employers, industry trainers, and educators, among others, Forman & Steen (1999) similarly argued that quantitative skills desired by employers are much broader than mere facility with the mechanics of addition, subtraction, multiplication, and division and familiarity with basic number facts; they also include some knowledge of statistics, probability, mental computation strategies, some grasp of proportional reasoning or modeling relationships, and broad problem-solving and communication skills about quantitative issues. Buckingham (1997), who studied what she called "specific and generic numeracies of the workplace" in some manufacturing industries in Australia, concluded that workplace numeracy is now about making decisions in the face of uncertainty in real situations, and that it encompasses far more than the basic skills traditionally associated with the term numeracy (as this term had been used in Australia).

Outcomes from skills projects are echoed in educational specifications. For example, basic skills projects in the United Kingdom and Australia influenced vocational education frameworks that name numeracy as an important skill, and describe stages or levels of accomplishment (Australian Education Council, 1992). The National Council for Vocational Qualifications Core Skills in the United Kingdom identified five levels of numeracy skill (Oates, 1992) that are closely linked to the sequence of content in the national school mathematics curriculum.

1.2 Broader life purposes

Since numeracy involves action in the real world, it is important to reflect on the kinds of purposes served by people's ability to act in a numerate way. Since people's numeracy is related to and may at times depend on people's literacy skills or other lifeskills, the purposes served by numeracy are expected to parallel those served by adults' literacy.

Work to describe the purposes served by adults' literacy and numeracy skills has been conducted in several countries. In Australia, for example, Kindler et al., (1996) reported on four such purposes: literacy for self-expression, literacy for practical purposes, literacy for knowledge, and literacy for public debate. In the Equipped for the Future initiative, The National Institute for Literacy in the United States has sponsored efforts to define critical skill areas. As part of the project, adult learners were asked what they needed to compete in a global economy and exercise the rights and responsibilities of citizenship. Content analysis yielded four broad types of purposes (Stein, 1995):

- Literacy for access and orientation in the world,
- Literacy as voice to one's ideas and opinions,
- Literacy for independent action, solving problems and making decisions as a parent, citizen and worker,
- Literacy as a bridge to further learning and to keep up with a rapidly changing world.

In Australia, a range of work has been done to create standards and a hierarchy of numeracy skill development that is not based upon school mathematics descriptions (Coates et al. 1995). In one key project (Kindler et al. 1996), numeracy was organized into four broad categories, according to different purposes and functions of using mathematics. *Numeracy for Practical Purposes* addresses aspects of the physical world that involve designing, making, and measuring. *Numeracy for Interpreting Society* relates to interpreting and reflecting on numerical and graphical information in public documents and texts. *Numeracy for Personal Organization* focuses on the numeracy requirements for personal organizational matters involving money, time and travel. *Numeracy for Knowledge* describes the mathematical skills needed for further study in mathematics, or other subjects with mathematical underpinnings and/or assumptions.

Overall, the purposes regarding literacy and numeracy appear to agree and suggest that adults need to be able to apply their numeracy and literacy skills to tasks with a social purpose in both informal and more formal contexts.

1.3 Educational perspectives

Recent years saw a growing dialogue about the goals and impact of mathematics education in schools. Various arguments have been brought forward to support a broadening of the conceptions regarding the mathematical skills and knowledge that school graduates should possess. In a society in which the media constantly present information in numerical or graphical form, the ability to interpret quantitative and statistical messages has been positioned by key stakeholders in education as vital for all adults (Steen, 1997). While employers have focused mostly on practical or job-specific numeracy skills, educators associated with the mathematical sciences have also paid much attention to the importance of quantitative literacy in civic and social contexts, and argued that mathematics is a crucial part of a common fabric of communication indispensable for modern civilized society, in part because it is the language of science and technology. Thus, understanding of public discussions and reports about socially important topics such as health and environmental issues are impossible without using the language of mathematics (National Research Council, 1989).

More mathematics educators now encourage links between the knowledge and skills gained in the mathematics classroom and students' ability to handle real-life situations that require activation of mathematical knowledge and skills (National Council of Teachers of Mathematics (NCTM),1989; Willis, 1990; Heuvel-Panhuizen & Gravemeijer, 1991). Such "handling" should be interpreted broadly, to mean not only application of mathematical procedures and concepts, but also many other abilities, such as the ability to critically reflect on information encountered (Frankenstein, 1989) or to understand and appreciate mathematical phenomena in the world, such as symmetry in the arts and nature.

The Realistic Mathematics Education (RME) initiative in the Netherlands that started in the early 1980s is an example of an attempt to develop educational experiences in light of the characteristics of real-world mathematical tasks and practices. RME is based upon the assumption that mathematics is an essential and important aspect of society, and therefore that mathematics education should be derived from real-life situations and should aim to create those skills applicable in any societal situation (family, work, etc). RME in adult education aims to optimize mathematical knowledge, skills, and problem-solving strategies that people have already been using in everyday life, or learned in or out of school, so that they can apply those strategies flexibly in all kinds of situations, have more control over their own personal, societal and work lives, and undertake further learning or training. Therefore, RME implies that school students and adults are to also be trained in cooperative learning and in recognizing and facilitating their own problem-solving procedures, strategies, and learning processes (van Groenestijn, 1998).

In the adult education sector, which is growing and becoming more formalized in many countries due to economic considerations and the need for lifelong learning, attention to mathematical skills is very visible (Benn, 1997). Educators working with adults aim to assist learners in developing mathematical concepts and relationships in ways that are personally meaningful but also functional. Adult educators usually assume that there is rarely only one right way, but a wide variety of strategies that work well when solving functional computational problems. Adults' personal methods of using mathematics are encouraged and valued. This is often a significant difference from traditional (pre-reform) school-based mathematics teaching, within which school students were often expected to follow the one correct method, or algorithm, introduced by the teacher to solve a problem.

The National Institute for Literacy in the United States has sponsored several efforts to define critical skill areas, as part of its Equipped for the Future initiative. One key project, by the Adult Numeracy Network (ANN) (Curry, Schmitt, and Waldron, 1996), was designed to reach a consensus on the kinds of mathematics that adults should know and hence are important to teach and assess in adult education. This project aimed to consolidate several curricular perspectives, mainly those offered by the NCTM (1989) the SCANS Commission (1991), and prior work by the ABE Mathematics Team in Massachusetts (Leonelli, Merson, Schmitt, and Schwendeman, 1994), as well as the results of interviews with hundreds of adult learners, numeracy teachers, and employers.

The ANN's *Framework for Adult Numeracy Standards: The Mathematical Skills and Abilities Adults Need to be Equipped for the Future*, organized needed knowledge into seven broad themes or areas: Relevance/connections, Problem solving/reasoning/decision making, Communication, Number and number sense, Data, Geometry: spatial sense and measurement, and Algebra: patterns and functions. The first three themes are concerned with processes of being numerate, while the latter four cover key content areas of mathematics. Again, this framework highlights and supports the view that numeracy is about making meaning of mathematical information, and that it encompasses a broad spectrum of skills and knowledge bases.

1.4 Research perspectives

A sizeable literature has accumulated over the last several decades regarding the ways in which adults use mathematical skills or cope with mathematical tasks in both formal (i.e., school-based) and informal (i.e., everyday or workplace) contexts. However, few attempts have been made to synthesize this literature and examine its implications for large-scale skills assessments. Some examples of research reports and theoretical discussions that could be considered in this regard include: Rogoff & Lave, 1984; Resnick, 1987; Saxe, 1988; Carraher, Schliemann, & Carraher, 1988; Scribner & Sachs, 1991; Nunes, 1992; or Coben, 2000.

Based on an analysis of the above and related literature, Gal (1993), while at the Numeracy Project at the National Center for Adult Literacy in the U.S., developed a conceptual perspective on the nature of adults' numeracy and numerate behavior. Gal (1993; 1997) argued that "numeracy" refers to the aggregate of skills, knowledge, and dispositions that enable and support independent and effective management of diverse types of quantitative situations. Gal further argued that the scope of adult numeracy is broad in light of the need for different types of responses in different situations.

Some situations call for generative responses, i.e., computing a number or generating an estimate or a decision. Examples are dealing with simple operations (measuring the length of a shelf), dealing with multi-step operations embedded in text (such as completing a tax form) and making reasonable decisions (for example, choosing the best loan). Other situations call for interpretive responses, i.e., making sense of quantitative statements or data displays (as in a newspaper article reporting crime statistics), and being able to ask critical questions about the information and arguments presented without performing any calculations. Both types of situations, and many mixed types, vary in terms of the literacy and communication skills they involve; in some cases it may not be possible to separate literacy from numeracy skills. It has also been suggested (Gal, 1997) that numerate behavior is enabled by *dispositional elements* (beliefs, attitudes, habits of mind) that motivate and support effective behavior in any given situation.

Gal (1993) also proposed that adults *manage* situations that call for application of numeracy skills. A person may decide to sacrifice precision or accuracy to reduce mental load or save time. A response may be reached in a computationally inefficient way or be based on non-standard procedures, but this may not matter in real-life as long as the individual expends time and effort in a way that is reasonable in light of the demands of the situation and his or her goals. It follows that there may be important differences between how adults respond to a school-oriented task (where adults may try to apply only school-based, memorized procedures), and demonstrate numerate, confident behavior in realistic situations. Cumming, Gal, & Ginsburg (1998) have argued that many of these aspects of numerate behavior are not reflected in how tests and test items are created and interpreted.

1.5 Assessment schemes

Some understanding of the mathematical needs of adult life can also be gleaned from an examination of large-scale assessment efforts, used either with adults or school students.

Adult assessments. A framework developed by Kirsch and Mosenthal (see Kirsch, Jungblut, & Mosenthal, 1998) to describe adults' literacy skills, including aspects of adult's quantitative skills, has been widely implemented in multiple national and international assessment projects, most recently the International Adult Literacy Survey (IALS; see Statistics Canada and OECD, 1996, 1997). The IALS framework made use of three literacy scales—Prose Literacy, Document Literacy, and Quantitative Literacy—to operationalize its conception of literacy. The ALL domain of numeracy is most closely related to the Document Literacy (DL) and Quantitative Literacy (QL) scales, which were defined as follows.

DL: *The knowledge and skills required to locate and use information contained in various formats (including job applications, payroll forms, transportation schedules, maps, tables, and graphics)*.

QL: *The knowledge and skills required to apply arithmetic operations, either alone or sequentially, to numbers embedded in printed materials (such as balancing a check book, figuring out a tip, completing an order form, or determining the amount of interest on a loan)*.

QL tasks as well as some DL tasks have addressed important aspects of people's mathematical knowledge and skills. For example, DL tasks required respondents to identify, understand, and interpret information given in various lists, tables, charts and displays; this information sometimes included quantitative information, such as numbers or percents. QL tasks required respondents to apply arithmetical operations learned mostly in elementary grades. However, these tasks did not require respondents to cope with other types of mathematical information (e.g., measurements, shapes) or with information whose processing does not require comprehension of text. In addition, tasks used in both scales called for a limited range of responses, i.e., exact computations or specific types of interpretations. Thus, while such tasks and responses are important by themselves, they represent only a subset of the much wider range of tasks and responses that are typical of many everyday and work tasks, such as sorting, measuring, estimating, conjecturing, or using models (e.g., formulas).

School assessments. IALS' central mission was assessing facets of real-world literacy, hence the QL scale focused on application of basic mathematical operations in response to functional tasks using realistic texts. Large-scale assessments of mathematical skills aimed at younger populations usually take quite different approaches. Selected such assessments, associated with the ongoing PISA (Project for International Student Assessment), organized by the OECD, and with the GED test in the US, are reviewed below. These assessments are reviewed to highlight issues that can inform the content of an assessment of adult numeracy skills, as well as to shed light on areas where an assessment of adult numeracy has to deviate from familiar forms of assessments that are common with school-age populations.

The PISA survey has developed a framework for assessing Mathematical Literacy, defined as follows:

> *An individual's ability to identify, to understand, to make well-founded judgments about, and to act towards the roles that mathematics plays in dealing with the world, as needed for that individual's current and future life as a constructive, concerned, and reflective citizen (PISA Mathematics Functional Expert Group, 1998).*

This definition shows some overlap and consistency with the conception of numeracy used in the present framework as well as with broader conceptions of literacy as adopted by IALS and ALL. Yet, some key differences seem to include the following:

- PISA focuses on how students understand, use, and apply mathematical skills and mathematize problems that are related to the *formal* school mathematics curriculum the students were expected to cover as part of their studies.
- PISA puts only partial emphasis on the realism of tasks. Given that students have limited world experience, tasks can be contrived or use formal symbolism that assesses mostly formal knowledge of what was taught in schools.

- The PISA mathematical assessment is not explicitly interested either in tasks where mathematical information is embedded in text (realistic or otherwise), or in the influence of literacy skills on mathematical performance (despite the inclusion of the term "literacy" in "mathematical Literacy").

Nationally-recognized standardized tests used in several countries to assess the mathematical knowledge of *adults* are often in line with school-based assessments. For example, the GED test in the U.S. (used to grant a high-school equivalency diploma to adults who did not formally graduate from high-school), and the National Vocational Qualifications system in the U.K., both use items with characteristics that are more in line with school-related assessments of mathematical knowledge than with the QL scale. These tests rely heavily on multiple-choice questions, employ some tasks requiring manipulation of numbers without a meaningful context, and require the use of some formal mathematical notations in formulas, either memorized or provided as part of the test.

School-oriented assessments point to some general areas of mathematical knowledge and skill that both school graduates as well as early school leavers may need to have to effectively cope with the various challenges of adult life. Reviewing the PISA, the *Third International Mathematics and Science Study* (TIMSS), and similar assessments highlights the fact that some important mathematical skills and knowledge that these assessments aim to capture were not captured by the QL scale of IALS. For example, knowledge of "big ideas" related to shape and geometry or to chance and statistics, knowledge of measurement systems, or the ability to "model" the mathematical aspects of certain situations were not included.

The above discussion is not meant to be a comprehensive review of current large-scale assessments of schooling-related mathematical skills (see Robitaille & Travers, 1992), nor a criticism of the QL scale of the IALS assessment framework. It simply reiterates that *all* assessments make conscious decisions regarding the (mathematical) skills that are important to assess, and that consequently the *forms* of assessment chosen carry not only advantages, but also disadvantages. The philosophy behind the design of mathematical assessments for PISA, GED, and similar assessments is based on assumptions about what it means to "know math" or "be able to do math" in a *schooling* context; hence, the assessment design assumes that it is legitimate to use a certain degree of formalization of math symbols and to present contrived math tasks. This assumption does not fit the assessment of skills of adults who may have been out of school for many years.

2. Towards a definition of Numeracy for ALL

The discussion above implies that adult numeracy should be viewed as different from "knowing school mathematics", that it is broader than the construct of Quantitative Literacy as defined in IALS, and also that multiple factors affect the way adults cope with the demands imposed by tasks that contain mathematical elements. (Note: For convenience, this report uses the term "mathematical" as inclusive of situations where *statistical* or *probabilistic* information may appear or where statistical thinking is required as well, even though statistical reasoning is not usually viewed as a branch of mathematics.).

Although a universally accepted definition of "numeracy" does not exist (Baker & Street, 1994), it is instructive to further examine some definitions and perspectives on the meaning of numeracy, and note that all contain an emphasis on the practical or functional application and use of mathematical knowledge and skills. The Australian Beazley Committee definition is typical:

> *Numeracy is the mathematics for effective functioning in one's group and community, and the capacity to use these skills to further one's own development and of one's community (Beazley, 1984).*

Another important element in defining numeracy is that of the role of communication processes. Numeracy not only incorporates the individual's abilities to use and apply mathematical skills efficiently and critically, but also requires the person to be able to interpret textual or symbolic messages as well as communicate about mathematical information and reasoning processes (Marr & Tout, 1997; Gal, 1997).

Most recent definitions of numeracy explicitly state that numeracy does not only refer to operating with numbers, as the word can suggest, but covers a wide range of mathematical skills and understandings. In recent years there has been much discussion and debate about the relationship between mathematics and numeracy and about the concept of "critical" numeracy (Frankenstein, 1989, Johnston, 1994). Johnston, for example, has argued that:

> *To be numerate is more than being able to manipulate numbers, or even being able to 'succeed' in school or university mathematics. Numeracy is a critical awareness which builds bridges between mathematics and the real-world, with all its diversity (Johnston, 1994).*

Many authors also argue that a discussion of functional skills should also address supporting or enabling attitudes and beliefs. In the area of adults' mathematical skills, "at homeness" with numbers or "confidence" with mathematical skills is expected, as these affect how skills and knowledge are actually put into practice (Cockroft, 1982; Tobias, 1993).

A definition of numeracy that seems to incorporate several of the aspects of numeracy noted above is from the Queensland Department of Education (1994) in Australia:

> *Numeracy involves abilities that include interpreting, applying and communicating mathematical information in commonly encountered situations to enable full, critical and effective participation in a wide range of life roles.*

An important commonality in the above descriptions of numeracy is the presence of mathematical elements in real situations, and the notion that these can be used or addressed by a person in a goal-oriented way, dependent on the needs and interests of the individual within the given context (home, community, workplace, etc.), as well as on his or her dispositions. Our earlier discussions further imply that numeracy involves more than just arithmetical skills as assessed in IALS, but extend to possession of number sense, estimation skills, measurement, and to multiple ways of responding flexibly to a mathematical situation. Finally, given the extent to which numeracy pervades the modern world, it is not necessarily just commonly encountered situations that require numerate behavior, but also *new* situations.

With the above in mind, a brief definition of numeracy proposed for the ALL is:

Numeracy: *The knowledge and skills required to effectively manage and respond to the mathematical demands of diverse situations.*

This brief definition of numeracy, and the fuller definition of "numerate behavior" provided in the next subsection, are much broader than the definition of QL as used in the IALS (see section 1.5). Its key concepts relate in a broad way to situation management and to a range of effective responses (not only to application of arithmetical skills). It refers to a wide range of skills and knowledge (not only to computational operations) and to a wide range of situations that present actors with mathematical information of different types (not only those involving *numbers* embedded in *printed* materials).

The brief definition above implies that numerate individuals are those who respond autonomously to situations in which mathematical ideas are embedded, actively using the power of mathematics rather than delegating or ignoring quantitative issues. However, a broad, inclusive definition alone is not enough to fully describe what a numerate person can do. The next section elaborates on specific facets of numerate behavior that underlie the general terms used in the above definition. These facets provide the detail necessary to develop an assessment tool to evaluate numeracy in adult populations.

3. Facets of Numerate Behavior

To develop a guide for the production of items, the main challenge was to determine how to assess, within the constraints of the ALL Survey protocol, the extent to which an adult effectively manages and responds to the mathematical demands of diverse situations. The numeracy team had to find ways to bound the range of tasks for assessing "numeracy" in a way that captures the breadth of the concept and yet is pragmatic for a large-scale assessment. Focusing on the idea of measurable numerate behavior was key to solving the challenge.

In actuality, people's numeracy is revealed through the responses or behaviors they generate (i.e., identifying, interpreting, acting upon, communicating) in reaction to the mathematical information or ideas that may be represented in a situation. As implied by the literature and ideas reviewed earlier in Section 1, the nature of a person's responses to the mathematical and other demands of a situation will depend critically on the activation of various enabling knowledge bases (understanding of the context; knowledge and skills in the areas of mathematics, statistics, and literacy), on reasoning processes, and on certain dispositions. It is clear that numerate behavior will involve an attempt to engage with a task and not delegate it to others or deal with it by intentionally ignoring its mathematical content.

Thus, a definition of numerate behavior that underlies the assessment of Numeracy in the ALL Survey is:

> ***Numerate behavior*** *is observed when people manage a situation or solve a problem in a real context; it involves responding to information about mathematical ideas that may be represented in a range of ways; it requires the activation of a range of enabling knowledge, factors, and processes.*

Table 1 lists specific components of five key facets of numerate behavior. These facets and their components are further explained in subsequent sections.

3.1 Facet 1: Contexts

People try to manage or respond to a numeracy situation because they want to satisfy a purpose or reach a goal. Four types of purposes and goals are described below. To be sure, these are not mutually exclusive and may involve the same underlying mathematical themes.

Everyday life. The numeracy tasks that occur in everyday situations are often management tasks that one faces in personal and family life. Others revolve around hobbies, personal development, and interests. Representative tasks are handling money and budgets, comparison shopping, personal time management, making decisions involving travel, planning holidays, mathematics involved in hobbies like quilting or wood-working, playing games of chance, understanding sports scoring and statistics, reading maps, and using measurements in home situations such as cooking or home repairs.

Table 1: Numerate behavior and its five facets

Numerate behavior involves:

(1) **managing a situation or solving a problem in a real context:**
everyday life
work
societal
further learning

(2) **by responding:**
identifying or locating
acting upon:
- order/sort
- count
- estimate
- compute
- measure
- model

interpreting
communicating about

(3) **to information about mathematical ideas:**
quantity & number
dimension & shape
pattern and relationships
data & chance
change

(4) **that is represented in a range of ways:**
objects & pictures
numbers & symbols
formulae
diagrams & maps
graphs
tables
texts

(5) **and requires activation of a range of enabling knowledge, factors, and processes:**
mathematical knowledge and understanding
mathematical problem-solving skills
literacy skills
beliefs and attitudes.

Work-related. At work, one is confronted with quantitative situations that often are more specialized than those seen in everyday life. In this context, people may develop good skills in managing situations that might be narrower in their application of mathematical themes. Representative tasks are completing purchase orders, totaling receipts, calculating change, managing schedules, budgets, and project resources, using spreadsheets, organizing and packing different shaped goods, completing and interpreting control charts, making and recording measurements, reading blueprints, tracking expenditures, predicting costs, and applying formulas.

Societal or community. Adults need to know about trends and processes happening in the world around them (e.g., regarding crime, health issues, wages, pollution) and may have to take part in social events or community action. This requires that adults can read and interpret quantitative information presented in the media, including statistical messages and graphs. Also, they may have to manage situations like organizing a fund-raiser, realizing the fiscal effect of community programs, or interpreting the results of a study of the latest health fad.

Further learning. It is often also important to have numeracy skills that enable a person to participate in further study, whether for academic purposes or as part of vocational training. In either case, it is important to be able to know some of the more formal aspects of mathematics that involve symbols, rules, and formulas and to understand some of the conventions used to apply mathematical rules and principles.

3.2 Facet 2: Responses

In different types of real-life situations, people may have to respond in one or more of the following ways (the first virtually always occurs; others will depend on the interaction between situational demands and the goals, skills, dispositions, and prior learning of the person):

Identify or locate some mathematical information present in the task or situation confronting them that is relevant to their purpose or goal.

Act upon or react to the information in the situation. Bishop (1988), for example, proposed that there are six modes of mathematical actions that are common in all cultures: counting, locating, measuring, designing, playing and explaining. Other types of actions or reactions may occur, such as doing some calculations (“in the head” or with a calculator), ordering or sorting, estimating, or modeling (such as by using or developing a formula).

Interpret the information embedded within the situation (and the results of any prior action) and comprehend what it means or implies. This can include making a judgment about how mathematical information or known facts actually apply to the situation or context. Contextual judgment may have to be used in deciding whether an answer makes sense or not in the given context, for example, that a result of “2.35 cars” is not a valid solution to how many cars are needed to transport a group. It can also incorporate a critical aspect, where a person questions the purpose of the task, the validity of the data or information presented, and the meaning and implications of the results, both for them as an individual and possibly for the wider community.

Communicate about the mathematical information given, or the results of one’s actions or interpretations to someone else. This can be done orally or in writing (ranging from a simple number or word to a detailed explanation or analysis) and/or through drawing (a diagram, map, graph).

3.3 Facet 3: Mathematical information

Mathematical information can be classified in a number of ways and on different levels of abstraction. One approach is to refer to fundamental "big ideas" in the mathematical world. Steen (1990), for example, identified six broad categories pertaining to: *Quantity, Dimension, Pattern, Shape, Uncertainty*, and *Change*. Rutherford & Ahlgren (1990) described networks of related ideas: *Numbers, Shapes, Uncertainty, Summarizing data, Sampling,* and *Reasoning*. Dossey (1997) categorized the mathematical behaviors of quantitative literacy as: *Data representation and interpretation, Number and operation sense, Measurement, Variables and relations, Geometric shapes and spatial visualization,* and *Chance*. The ALL Numeracy team drew from these three closely tied categorizations to arrive at a set of five fundamental ideas that in their view characterize the mathematical demands met by adults in diverse situations at the beginning of the 21st century.

Quantity and Number. *Quantity* is described by Fey (1990) as an outgrowth of people's need to quantify the world around us, using attributes such as: length, area, and volume of rivers or land masses; temperature, humidity, and pressure of our atmosphere; populations and growth rates of species; motions of tides; revenues or profits of companies, etc. *Number* is fundamental to quantification and different types of number constrain quantification in various ways: whole numbers can serve as counters or estimators; fractions, decimals and percents as expressions of greater precision, parts or comparisons (ratios); and positive and negative numbers as directional indicators. In addition to quantification, numbers are used to put things in order and as identifiers (e.g., telephone numbers or zip codes). Facility with quantity, number, and operation on number requires a good "sense" of magnitude. Contextual judgment comes into play when deciding how precise one should be or which tool (calculator, mental math, a computer) to use. Money and time management, the ubiquitous mathematics that is part of every adult's life, depends on a good sense of number and quantity. A basic level numeracy task might be figuring out the cost of one can of soup, given the cost of 4 for $2.00; a task with a higher cognitive demand could involve more complex numbers such as when figuring out the cost per pound when buying 0.783 kg of cheese for 12,95 Euros.

Dimension and shape. *Dimension* includes "big ideas" related to one, two, and three dimensions of "things" (using spatial and numerical descriptions), projections, lengths, perimeters, planes, surfaces, location, etc. Facility with each dimension requires a sense of "benchmarks" and estimation, direct measurement and derived measurement skills. *Shape* is a category describing real images and entities that can be visualized (e.g., houses and buildings, designs in art and craft, safety signs, packaging, snowflakes, knots, crystals, shadows and plants), as well as highly abstract "things" greater than three dimensions. Direction and location are fundamental qualities called upon when reading or sketching maps and diagrams. A basic numeracy task in this fundamental aspect could be shape identification whereas a complex task might involve describing the change in the size of an object when one dimension is changed.

Pattern, Functions and relationships. It is frequently written that mathematics is the study of patterns and relationships. Pattern is seen as a wide-ranging concept that covers patterns encountered all around us, such as those in musical forms, nature, traffic patterns, etc. It is argued by Senechal (1990) that our ability to recognize, interpret, and create patterns is the key to dealing with the world around us. The human capacity for identifying relationships and for thinking analytically undergirds mathematical thinking. Algebra - beyond symbolic manipulation - provides a tool for representing relationships between amounts through the use of tables, graphs, symbols, and words. The ability to generalize and to characterize functions, relationships between variables, is a crucial gateway to understanding even the most basic economic, political or social analyses. A basic level numeracy task might require someone to describe how items are arranged in a package; developing a formula for an electronic spreadsheet would put a higher level of demand on the individual.

Data and chance. Data and chance encompass two related but separate topics. *Data* covers "big ideas" such as variability, sampling, error, or prediction, and related statistical topics such as data collection, data displays, and graphs. Modern society demands that adults interpret and produce organizers of data such as frequency tables, pie charts, graphs and to sort out relevant from irrelevant data. *Chance* covers "big ideas" related to probability, subjective probability, and relevant statistical methods. Few things in the world are 100% certain; thus the ability to attach a number that represents the likelihood of an instance is a valuable tool whether it has to do with the weather, the stock-market, or the decision to board a plane. In this mathematical category, a simple numeracy skill might be the interpretation of a simple pie chart; a more complex task would be to infer the likelihood of an occurrence, such as predicting the weather, based upon past information.

Change. This term describes the mathematics of how the world changes around us. Individual organisms grow, populations vary, prices fluctuate, objects traveling speed up and slow down. Change and rates of change help provide a narration of the world as time marches on. Additive, multiplicative, exponential patterns of change can characterize steady trends; periodic changes suggest cycles and irregular change patterns connect with chaos theory. Describing weight loss compares as a simple task to calculating compounded interest.

3.4 Facet 4: Representations of mathematical information

Mathematical information in an activity or a situation may be available or represented in many forms. It may appear as *concrete objects* to be counted (e.g., people, buildings, cars, etc.) or as *pictures* of such things. It may be conveyed through symbolic notation (e.g., numerals, letters, and operation or relationship signs). Sometimes, mathematical information will be conveyed by *formulas*, which are a model of relationships between entities or variables.

Mathematical information may be encoded in visual displays such as a *diagram* or *chart*; *graphs* and *tables* may be used to display aggregate statistical or quantitative information (by displaying objects, counting data, etc.). Similarly, a *map* of a real entity (e.g., of a city or a project plan) may contain information that can be quantified or mathematized.

Finally, a person may have to extract mathematical information from various types of *texts*, either in prose or in documents with specific formats (such as in tax forms). Two different kinds of text may be encountered in functional numeracy tasks. The first involves mathematical information represented in textual form, i.e., with words or phrases that carry mathematical meaning. Examples are the use of number words (e.g., "five" instead of "5), basic mathematical terms (e.g., fraction, multiplication, percent, average, proportion), or more complex phrases (e.g., "crime rate increased by half") which require interpretation. The second involves cases where mathematical information is expressed in regular notations or symbols (e.g., numbers, plus or minus signs, symbols for units of measure, etc.), but is surrounded by text that despite its non-mathematical nature also has to be interpreted in order to provide additional information and context. An example is a bank deposit slip with some text and instructions in which numbers describing monetary amounts are embedded.

3.5 Facet 5: Other enabling factors and processes

The way in which each person manages his interpretations and responses to the contexts, tasks, and mathematical representations described above as facets of adult numeracy depends of course on his or her mathematical knowledge, whether formally learned, informally developed, or self-invented. This includes but is not limited to the understanding and ability to apply concepts, ideas and procedures detailed in many strands for school curricula or in many school-based assessments under titles such as Whole numbers and basic operations, Ratios, percents, decimals, and fractions, Measurement, Geometry, Algebra, or Probability and statistics. These topics are interwoven into the five areas of "mathematical information" described above in subsection 3.3, and are assessed by the items in the Numeracy scale.

Numerate behavior, however, depends on an integration of mathematical knowledge bases with broader reasoning and problem-solving skills and strategies needed to be able to think and to act mathematically. Further, numerate behavior depends on the integration of the above with the literacy skills, the dispositions (beliefs, attitudes, habits of minds, etc), and prior experiences and practices that an adult brings to each situation. These are briefly discussed below, with comments on the extent to which each is assessed in ALL.

Problem-solving skills. Throughout life, adults develop or apply diverse strategies to manage their quantitative situations. Some strategies or skills may be based on prior formal learning, while others may be self-invented or adapted to fit the situation at hand. To solve many computational problems or to figure a way to manage certain quantitative tasks, people have to re-construct reality in a mathematical way, for example, model or mathematize. They can do so either on their own or in discussion with other people. Problem-solving strategies may include, e.g., extracting relevant information from the task/activity; rewriting/restating the task; drawing pictures, diagrams or sketches; guessing and checking; making a table; and/or generating a concrete model or representation. To some extent, these strategies are appropriate for determining a response to the Numeracy items, but they are assessed more fully in the Problem Solving Scale of the ALL Survey.

Literacy skills. The ability to read, write, and talk are important skills in undertaking a numeracy task or activity or communicating the outcomes of working on such tasks. In cases where “mathematical representations” involve text, one's performance on numeracy tasks will depend not only on formal mathematical or statistical knowledge but also on reading comprehension and literacy skills, reading strategies, and prior literacy experiences. For example, following a computational procedure described in text (such as the instructions for computing shipping charges or adding taxes on an order form) may require special reading strategies, as text is very concise and structured. Likewise, analyzing the mathematical relationships described in words requires specific interpretive skills, e.g., realizing that “four more than” is a different relationship than “four times as much.”

Beliefs and attitudes. A person may not necessarily act in numerate ways, even if she or he can demonstrate high ability on a numeracy test. The way in which a person responds to a numeracy task—including overt behavior or actions as well as cognitive processes and the propensity to adopt a critical stance—will depend not only on knowledge and skills but also on beliefs, attitudes, habits of mind, and prior practices. In some cultures, some adults, including highly educated ones, decide that they are not "good with numbers." These sentiments or self-perceptions are usually attributed to negative prior experiences they have had as pupils of mathematics (Tobias, 1993), and stand in contrast to the desired sense of "at-homeness with numbers" (Cockcroft, 1982). Such attitudes and beliefs can interfere with one's motivation to develop new mathematical skills or to tackle math-related tasks, and may also affect test performance (McLeod, 1992).

In realistic contexts, adults with a negative mathematical self-concept may elect to avoid a problem with quantitative elements, address only a portion of it, or prefer to delegate a problem, e.g., by asking a family member or a salesperson for help. Such decisions or actions are indeed the prerogative of a manager and can serve to reduce both mental and emotional load (Gal, 2000). Yet, such actions may fall short of autonomous engagement with the mathematical demands of real-world tasks (as noted in the core definition of numeracy used here), carrying negative consequences, e.g., not being able to fully achieve one's goals. Therefore, prior experiences and existing habits of coping with mathematical and numeracy situations may be influential.

Numeracy-related practices and experiences. Research suggests that, for adults as well as for children, mathematical knowledge develops both in and out of school (e.g., Schliemann & Acioly, 1989; Saxe, 1991; Lave, 1998). Saxe and his colleagues have written about the importance of cultural practice in the development of mathematical thinking and how such practices profoundly influence an individual's cognitive constructions and mathematical ideas, depending, e.g., on the artifacts or tools they use, the nature of the measurement systems in their culture, the counting or calculating devices (abacus, calculator) they use, the distribution of work among family members, or general patterns and types of social activity.

Mathematical experiences and practices, whether at work, home, when shopping, or in other contexts, can be both the result of a certain skill level, or the cause of skill levels. For this reason, it was deemed important to add several items to the Background Questionnaire that examine the frequency of performing certain related numeracy tasks in different contexts including how often artifacts such as calculators or computer spreadsheets are used.

4. Factors affecting complexity of Numeracy Items

Because of the scarcity of research on adults' use and application of numeracy, there is insufficient empirical knowledge to determine what factors make a numeracy activity or task more difficult or complex. One of the more exciting, and challenging, aspects of the project was the development of a scheme to account for the difficulty of different numeracy assessment tasks. We sought such a scheme to inform item development, i.e., help in the creation of items that spread over a range of difficulty levels. However, if the scheme could be shown to correlate with actual difficulty levels of items as measured in actual testing of a sample of individuals, it could also be used to help *explain* observed performance. Given its importance for both item development and interpretation of results, the complexity scheme is described in detail in this section.

4.1 Previous Research on task complexity.
In IALS, three factors were found to be the principal components of task difficulty (regarding literacy or text-based tasks): plausibility of distractors, type of match required, and type of information required. The difficulty of the Quantitative Literacy tasks appeared to be a function of several other factors:

1. The particular arithmetic operation required to complete the task
2. The number of operations needed to perform the task
3. The extent to which the numbers are embedded in printed materials
4. The extent to which an inference must be made to identify the type of operation to be performed (i.e. problem transparency; see below)

The IALS QL difficulty factors overall fit those used in large-scale assessments of mathematical skills (with children), which often make use of three or four factors:

1. *The mathematical concepts involved:* number systems and number sense, spatial and geometrical topics, functions and algebra, chance/statistics topics, etc. Concepts that are related to topics taught in lower grades are considered easier.
2. *The complexity of operations*: addition, subtraction, multiplication, and division, as well as dealing with whole numbers, with decimals, and with percents. Operations that are related to topics taught in lower grades are considered easier.
3. *The number of operations*: one-step problems are considered easier than multi-step problems.
4. *Problem transparency*: This factor is sometimes relevant; it refers to the extent to which the problem situation includes clearly identified numbers or entities and the extent to which it is clear what operations or actions to perform. To the extent that these are not clear or transparent, respondents have to extract needed information by applying comprehension and inference strategies, making the task more complex.

There are other adult-related assessment projects on which to draw to develop the levels of complexity. Both the Essential Skills Research Project and the Applied Numeracy sub-test of the Work Keys test battery (American College Testing, 1997) use a two-factor model of complexity in their description of numeracy levels. The first factor, "operations required;" is seemingly straightforward and refers to the difficulty of operations called for. However, this is complicated by the level of difficulty of the numbers being manipulated: computations that include fractions and decimals are usually more difficult than those with whole numbers.

The Essential Skills model spells out two sequences of complexity on this factor: *Operations* and *Translation* of information (sometimes called 'problem transparency').

Operations:

1. Only the simplest operations are required and the operations to be used are clearly specified. Only one type of mathematical operation is used in the task.
2. Only relatively simple operations are required. The specific operations to be performed may not be clearly specified. Tasks involve one or two types of mathematical operation. Few steps of calculations are required.
3. Task may require a combination of operations or multiple-applications of a single operation. Several steps of calculation are required. (More advanced operations may call for multiplication or division.)
4. Tasks involve multiple steps of calculation.
5. Tasks involve multiple steps of calculation. Advanced mathematical techniques may be required (e.g., percents, ratios, proportions).

Translation (Problem Transparency)

1. Only minimal translation is required to turn the task into a mathematical operation. All the information required is provided.
2. Some translation may be required or the numbers needed for the solution may need to be collected from several sources. Simple formulae may be used.
3. Some translation is required but the problem is well defined.
4. Considerable translation is required.
5. Numbers needed for calculations may need to be derived or estimated; approximations may need to be created in cases of uncertainty and ambiguity. Complex formulae, equations or functions may be used.

Two considerations prompted us to question the appropriateness of using mathematics-related frameworks (from Essential Skills or elsewhere) as the sole source for development of a complexity scheme for items assessing *adults'* ability to cope with real-world numeracy tasks. First, effective coping with many real-world quantitative problems depends upon people's ability to make sense of and interact with different types of texts. This is hardly recognized by the Essential Skills model. Hence, it was essential to add difficulty factors that acknowledge the inherent links between literacy and numeracy, quite similar to those used in IALS.

Another, albeit a more restricted consideration, is that the ordering of complexity of tasks by the type of operation performed may not be as clear with adults as it may be with children. Such ordering in school-based assessments is predicated on traditional school curricula, where more advanced topics are learned at higher grades. However, adults are known to use a lot of invented strategies, perhaps more so, and more efficiently so, than children. Multiplication or division problems, which can prove relatively hard for some young people, may be solved by seemingly simpler strategies, such as by repeated addition or repeated subtraction; complex numbers may be broken down in ways that ease mental load, and so forth. In addition, adults' familiarity with everyday contexts, such as with monetary entities, facilitates their performance with some seemingly advanced concepts. For example, specific benchmark values of fractions and percents, such as 1/2, 1/4, 50%, or 25%, are familiar to many people; as a result, they may be easier to manage than expected, violating curriculum-based ordering of difficulty. Hence, an *overall complexity level* has to be used, in order to weight these "inconsistencies" in ordering of difficulty levels proposed in other schemes.

4.2 Complexity factors in the ALL survey

The above literature review suggests that a framework of factors affecting the complexity of numeracy tasks should not only address factors related to the numerical and textual aspects of tasks, but should also address other issues. It should treat separately the number of operations and the type of operations from the *type of mathematical (or statistical) information to be processed*, which may involve numbers explicitly but also other types of mathematical information. In so doing, the desired framework of complexity factors should take into account the broad scope of the definition of numeracy, i.e., reflect the variation within contexts, the range of mathematical ideas/content, the types of possible responses, and the types of representations that cut across adult life contexts.

With the above considerations in mind, five key factors have been identified that are predicted to affect, separately and in interaction, the difficulty level of numeracy tasks to be used in the ALL survey. These five "complexity factors" are outlined in Table 2 and are organized in two sets: two factors that address mainly textual aspects of tasks, and three factors that address the mathematical aspects of tasks. These five factors are listed separately for clarity of presentation, but in actuality are *not* independent of each other and do interact in complex ways. Each factor is examined in some detail below, followed by a later subsection that describes the calculation of an overall complexity level for each item, taking into account all five factors.

Table 2: Complexity Factors—Overview

Aspects	Category	Range
Textual aspects	1. Type of match/problem transparency	Obvious/explicit to embedded/hidden
	2. Plausibility of distractors	No distractors to several distractors
Mathematical aspects	3. Complexity of Mathematical information/data	Concrete/simple to abstract/complex
	4. Type of operation/skill	Simple to complex
	5. Expected number of operations	One to many

Type of Match/Problem Transparency. This is a combination of the factor of Problem Transparency outlined above, and of an IALS factor called Type of Match. Problem Transparency is a function of how well the mathematical information and tasks are specified and includes aspects such as how apparently the procedure is set out, how explicitly the values are stated, etc. Type of Match refers to the process that a respondent has to use to relate the requested action in the question to the information in the task or text, which can range from a simple action of locating or matching to more complex actions that require the respondent to perform a number of searches through the information given. This measure of complexity for a numeracy task incorporates the degree of text embeddedness of the mathematical information.

In easy tasks, the type of information (e.g., numerical values) and the operations needed are apparent and obvious from the way the situation is organized. In more difficult ones, the values must be located or derived from other values; the operations needed may have to be discovered by the performer, depending on his or her interpretation of the context and of the kind of response expected. As well, numeracy situations may involve text to varying degrees, and this text may be of different degrees of importance. There may be a situation where there is little or no text. Some situations may involve pure quantitative information that is to be interpreted or acted upon with virtually no text or linguistic input. In other words, the performer derives all the information needed to respond from the objects present in the situation or from direct numerical displays.

At a higher level, some textual or verbal information may be present alongside the mathematical information. The text can provide background information about the problem situation, or some instructions. For example, a bus schedule, cooking instructions, and a typical school-type word problem all involve some text and some numbers. Still other situations would be heavily text-based or may not involve any numbers or mathematical symbols at all, just plain text. The task will contain mathematical or statistical information that a person needs to understand and, in some cases, act upon, but it will be much less transparent. It may be heavily embedded in dense text or may require using information from a number of sources within or even outside the text/task.

This factor requires that a task will be analyzed in terms of the questions: *How difficult is it to identify and decide what action to take?*, and *How many literacy skills are required?*

Plausibility of Distractors. This variable is literacy related, even though it can involve mathematical components. In general, literacy tasks are easiest to process when there are no plausible distractors in the text, that is, there is no other information in the text that meets any of the requirements of the task. At higher levels of difficulty, tasks can involve irrelevant information both within the question as well as within the text. In terms of mathematical information, a low level of plausible distractors would mean that no other mathematical information was present apart from that requested, making the numbers or data required easy to identify. At a higher level, there may be either some other mathematical information in the task (or its text) that could be a distractor, or the mathematical information given or requested could occur in more than one place. A higher level of complexity could also mean that outside information (e.g. the knowledge of a formula) may be needed to answer the question.

This factor requires that a task will be analyzed in terms of the questions: *How many other pieces of mathematical information are present?*, and *Is all the necessary information there?*

Complexity of Mathematical Information. Some situations present a person with simple mathematical information, such as concrete objects (to be counted), simple whole numbers, or simple shapes or graphs. At lower skill levels, the information will be more familiar, whereas at higher levels, the information may be less familiar. Situations will be more difficult to manage if they involve more abstract or complex information, such as very large or very small numbers, unfamiliar decimals or percents, information about rates, or dense visual information, as in a diagram or complex table.

This factor requires that a task will be analyzed in terms of the question: *How complex is the mathematical information that needs to be manipulated or managed?*

Type of Operation/Skill. Some situations require simple operations, such as addition or subtraction, or simple measurement (e.g., finding the length of a shelf), or recognition of shape. These are usually easier to analyze mathematically than situations that require multiplication or division, and than situations that require using exponents. While the difficulty of recognizing and carrying out the operation implied by a situation (be it additive, multiplicative, etc.) has direct bearing on task complexity, there may be exceptions that occur when alternative approaches are obvious. There are some tasks that combine both interpretive and generative skills and may involve a deeper conceptual understanding than merely carrying out a procedure. Other more complex tasks may involve an explanation of one's reasoning. The interpretation of information appearing in graphs, for example, becomes more complex if comparisons, conjecturing, or "reading beyond the information given" is required.

This factor requires that a task will be analyzed in terms of the question: *How complex is the mathematical action that is required?*

Expected Number of Operations. Tasks that require acting upon the mathematical information given may call for one application (step) of an operation, or for one action (e.g., literal reading of information in a table, or measurement). More complex tasks will demand more than one operation, which may be the same or similar to one another, such as the steps involved in multiple passes on the data or text. Still more complex tasks are those that involve the integration of several different operations.

This factor requires that a task will be analyzed in terms of the question: *How many steps and types of steps are required?*

4.3 Overall Complexity Level

It is possible to estimate the overall difficulty level of a specific item by first scoring the item on each of the five factors of complexity, according to the levels described in Appendix 2, and then summing together the scores for each factor. Figure 1 below explains the process; Appendix 2 describes each level of the five factors in detail. The total summary score can range between 5 (easiest) and 19 (most difficult).

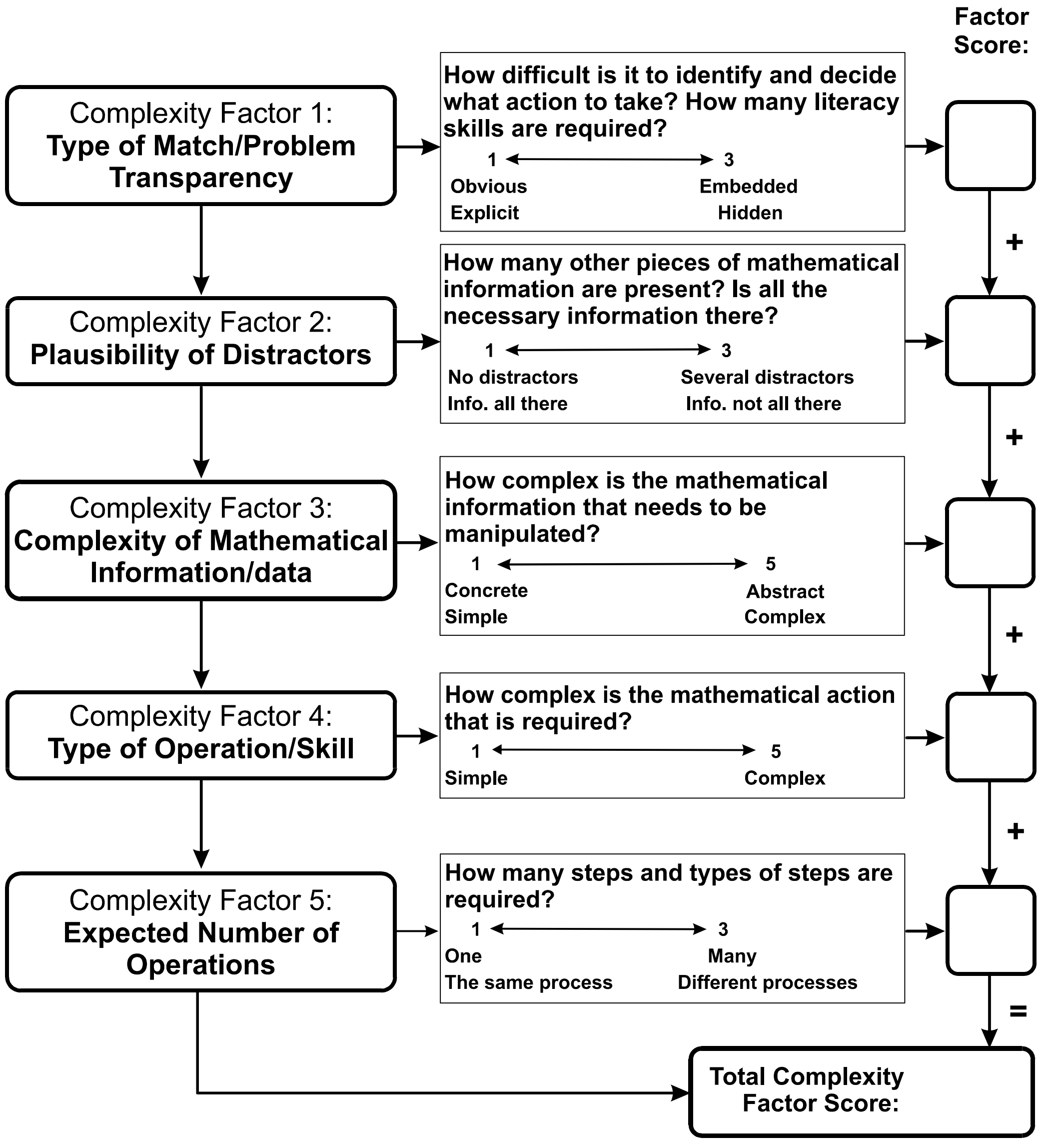
The estimation process outlined in Figure 1 suggests that each factor has a separate contribution to an item's overall difficulty or complexity. However, it can be hypothesized that as tasks become more complex, actual performance on items may increasingly depend not only on each factor by itself, but also on the interplay or interaction between them. Hence, the computational process suggested in Figure 1 can provide only approximate information about an item's anticipated difficulty level.

Further, the difficulty of a task cannot in some cases be predicted without taking into account characteristics of the person who interacts with the task. The same task may be more difficult for some individuals and less difficult for other individuals, depending on factors such as their familiarity with the context in which a task is situated, knowledge of formal mathematical notations, background world knowledge, as well as general literacy, problem-solving, and reasoning skills. For example, it could be predicted that a task that involves the composition of a fertilizer would be more difficult for an urban apartment dweller than for a rural farmer whereas a task that uses a bus schedule would be more difficult for the farmer. For the above reasons, the prediction of the difficulty of a task in isolation of detailed knowledge about the respondent himself can only be an estimate.

Despite the above limitations, the scheme of complexity factors developed for numeracy assessment in ALL comprises a theoretical contribution. It provides a conceptual basis for predicting the different levels of complexity of a broader range of items well beyond those involving arithmetic operations only. To the extent these predicted difficulty levels later prove to correlate with actual difficulty (such as in terms of percent correct on different items), this scheme can also assist in interpreting survey results, as it can help to explain *why* some items are harder than others and what kinds of cognitive processes are called for by different tasks.

Indeed, results from the feasibility and pilot studies described later in this report show that predicted difficulty and observed difficulty were highly correlated. Nonetheless, further distillation and validation are needed, and the scheme in its current state is included mainly to show the logic behind a process used during item development.

Figure 1. Complexity Flow chart

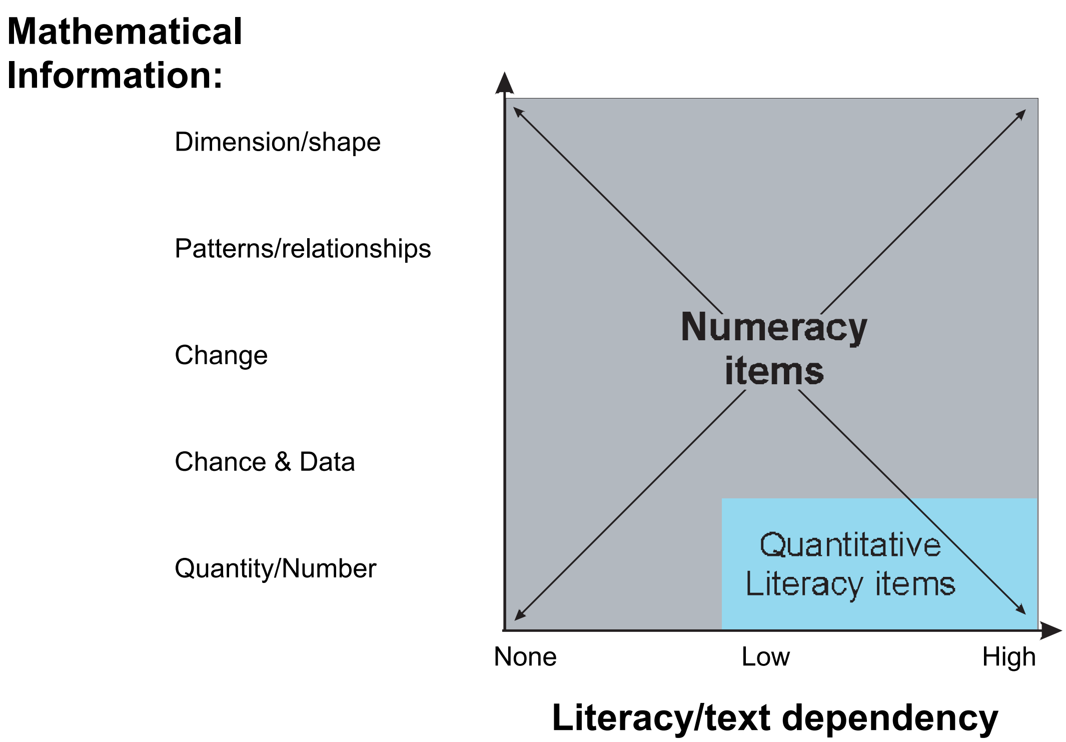


5. Summary: Directions and challenges in assessing Numeracy in ALL

Given the increasing need for adults to continuously adapt to changing citizenship and workplace demands (European Commission's White Paper, 1996), it is vital that nations have information about their workers' and citizens' numeracy in order to understand skill distributions in the population as a whole as well as in specific groups, and be able to plan effective lifelong learning opportunities.

Numeracy, a domain to be assessed in the ALL survey, has been conceptualized in this report as a broader construct than Quantitative Literacy that was assessed in the earlier IALS. Numeracy requires more varied responses (order, count, estimate, compute, measure, interpret, explain) to a wider range of mathematical information (quantity, dimension and shape, pattern and relationships, data and chance, and change) that may be embedded in text in varying degrees. Figure 2 illustrates the difference between the scope of topics that should be encompassed by Numeracy and Quantitative Literacy items.

Figure 2. Numeracy versus Quantitative Literacy



Ideally a survey of adult numeracy skills would ask respondents to complete tasks that are couched in real-life situations and that encompass all components of the facets of numerate behavior described earlier. Further, respondents would ideally respond to in-depth and qualitative interviews in order to understand their answers and reactions as well as their thought processes. Since the ALL is a paper-based household survey, however, there are restrictions on the 'real' nature of the situations that could be included. In addition, the available options for presenting questions and for eliciting, recording, and reliably scoring responses are limited when large samples across wide ranging populations have to be surveyed. Finally, given that numeracy involves a sizable number of facets and subcomponents, practical limitations on the number of items that can be administered to respondents to assess each of the domains covered in ALL mean that not all aspects of the facets of numerate behavior can necessarily be covered at an equal depth.

Operating within these constraints, the goal for the item development process described in the next part was to create Numeracy items that are based on realistic stimuli and include tasks that serve a purpose for most people within their culture. A scheme of factors affecting the complexity of numeracy tasks was developed as a tool to support item development and to help explain performance on Numeracy items. Information derived from performance on the Numeracy scale was designed to be supplemented by information gathered via the Background Questionnaire about prior practices and attitudes related to numeracy, in addition to other important explanatory variables such as those related to demographic data and literacy practices. The process of developing and evaluating items along with their scoring rubrics is described in Part B.

Closing this summary of challenges in converting the conceptual framework of Numeracy into an assessment tool, it must be noted that Numeracy cannot be fully separated from Document Literacy and Problem Solving, two other important domains assessed by ALL. Numeracy is somewhat related to Document Literacy, due to the role that text and graphical or tabular displays play in both constructs. Likewise, Numeracy is not fully distinguishable from Problem Solving, as certain numeracy tasks require the implementation of problem solving processes, albeit brief at times, while in contrast certain "problems" (e.g., designing a budget) involve numerical content.

Hence, a final challenge in designing items for the Numeracy scale was to achieve a sensible "division of labor". On one hand, it is important to prevent excessive overlap between what is measured by the Numeracy scale and the type of skills assessed by scales in other domains. On the other hand it is also important to maintain a reasonable coverage of the core issues encompassed by the construct of numeracy as envisioned in this report, even if this creates some acceptable degree of overlap with other scales. Such overlap, we argue, is not a limitation of the conceptual framework developed to describe Numeracy and numerate behavior, but simply reflects the complex nature of real-world situations and tasks. While it is convenient for educators and designers of assessment tools to think of skill areas as distinct and having clear-cut boundaries, real-life situations may require that people who have to manage and respond to them rely on interrelated knowledge and skills.

Part B: Item Development

6. Approach

Throughout the item-production stages, the Numeracy team used the principles or guidelines outlined below as the basis for creating items. (Note: Appendix 1 includes sample items illustrating the type of stimuli and questions developed).

a. Items should cover as many aspects within each of the four facets of the numeracy construct as possible. Items were generated so as to require the activation of a broad range of skills and knowledge included in the construct of numeracy, as portrayed in the conceptual framework described in Part A and depicted in Table 1.

b. Items should include a realistic stimulus and one or more questions that adults might be expected to encounter, given such a stimulus. The aim was for the item content to be familiar and the questions purposeful to respondents across cultures. Most tasks were derived from real-life stimuli, from situations that are easily understood or that can be expected to be of importance or relevance in different cultures.

c. Items should have a free-response format. Items were structured to include a stimulus (e.g., a picture, drawing, display) and one or more questions, the answers to which the respondent writes in his own words.

d. Items should spread over different levels of difficulty. Items were produced to span the range of difficulty levels covered by the theoretical complexity factors outlined earlier. Attention was paid to generating some items at the lowest level of complexity, which are of interest in countries where policies and educational programs may be earmarked for low-skill populations.

e. Items should vary in the degree to which the task is embedded in text. While about 1/3 of items were embedded in or included relatively rich texts, about 2/3 of the items were designed to use little or no text, to reduce overlap with the Document and Prose literacy scales as they existed in IALS and continue in ALL.

f. Tasks should be relatively short. The use of short tasks enabled the inclusion of a large number of diverse stimuli and questions, thus allowing for coverage of many key facets of the numeracy framework. Items were not designed to simulate extended problem-solving processes, thus reducing overlap with the Problem-Solving scale.

g. Items should overcome issues associated with different unit systems. Items were designed so that different currency systems as well as different systems of measurement (metric or Imperial) could be applied to the numbers or figures used. Such items can retain equivalency with respect to their mathematical or cognitive demands and appear familiar to different populations even after being translated.

7. Production and evaluation of items

The creation of items for the Numeracy assessment progressed through three stages: Two stages involving production of items and their testing in two countries on relatively small samples, and the third stage involved a much larger Pilot testing process.

7.1 Stage 1 (1998-1999): Production and field-testing of a first item pool.
Based on the above general principles, a pool of over 80 items was generated by team members, based on their experience in research, assessment, and teaching with both school-based and diverse adult and workplace learner populations in several countries.

Production grid. Items were created so as to fill cells within an item production grid with four key dimensions that match the conceptual facets outlined in Table 1:

1. Type of purpose / context: *everyday, societal, work, further learning*.
2. Type of response: *identifying or locating, acting upon* (*order/sort, count, estimate, compute, measure, model) interpreting, communicating about*.
3. Type of mathematical or statistical information: *quantity, dimension, patterns/relations, data/chance, change*. The content of the tasks was also conceived, however, in terms of common school-based mathematics topics more familiar to policy makers and educators, i.e., whole numbers and basic operations; ratios, percents, decimals and fractions; measurement; geometry; algebra; and statistics.
4. Type of representation of mathematical or statistical information: *numbers, formulae, pictures, diagrams, graphs, tables, texts*.

Scoring. Guidelines for scoring responses were designed to classify them into three general groups: "correct", "any other response" (i.e., wrong answers) and "not attempted" (i.e., no indication the respondent tried an item). However, for many items, *multiple* codes were prepared to capture different types of "correct" or "wrong" answers and thus enable an analysis of error patterns and shed light on the extent to which instructions are understood and items elicit the expected type of responses. In some items that require estimation or measurement, multiple codes were prepared to capture responses that may have different degrees of accuracy yet still fall within a "correct" or "wrong" region, in order to understand the level of accuracy that respondents adopt.

Non-cognitive items. Research literature suggests that the way in which a person responds to a numeracy task, including overt actions as well as internal thought processes and the adoption of a critical stance, depend not only on knowledge and skills but also on negative attitudes towards mathematics, beliefs about one's mathematical skills, habits of mind, and prior experiences involving tasks with mathematical content (Cockcroft, 1982; Lave, 1988; Schliemann & Acioly, 1989; Saxe, 1991; McLeod, 1992; Gal, 2000). Hence, the Numeracy team also prepared several scales for the Background Questionnaire, with questions designed to measure numeracy practices at home and at work, attitudes and beliefs about mathematics, and information about the environment in which the respondent learned mathematics while in school. Such scales may help in explaining performance on numeracy tasks, as well as understanding respondents' status on variables of interest to policy makers, such as participation in further learning or employment status.

First feasibility study. Of the items generated at this stage, 80 items were tested in a feasibility study conducted in the USA and The Netherlands, on samples of about N=300 per country; each item was answered by about 150 respondents. Results enabled analysis of error patterns and gender bias on items, and assessment of psychometric properties of items in terms of both classical test theory and Item Response Theory (IRT) parameters. Comments made by respondents in focus groups suggested that only a few items were at times misunderstood. Statistical analyses showed that most items have adequate psychometric properties, are answered roughly in the same way by males and females, and that the items tested cover a wide range of difficulty levels. (Four items from the QL scale used in IALS were also tested to enable rough calibration of the item difficulty estimates obtained on the basis of the feasibility sample, in light of the difficulty levels of these four items in the much larger and nationally representative samples used in IALS).

The feasibility study enabled evaluation of several other important issues that arise during item translation and adaptation to different cultural contexts, e.g., in terms of using different monetary, length, or volume units. Results showed that performance on most items is comparable across languages even when they use different units of measurement. Also, responses were scored independently by two scorers, and an analyses of scorer agreement showed very adequate scoring reliability, suggesting that the scoring schemes and scoring instructions were well understood to scorers operating in two languages.

Further analyses revealed interesting patterns of correlations between respondents scores on some of the non-cognitive scales and their overall performance on Numeracy items. This suggested that some of the attitudinal and belief items can indeed be used to help in understanding cognitive performance.

Overall, 68 items out of the 80 items tested appeared to satisfy all selection criteria. For these 68 items, it was found that item difficulty (in terms of percent correct on an item) was highly correlated ($r = -0.793$) with predicted item complexity as determined during item development based on the factors outlined in Figure 2 and detailed in Appendix 2. This indicated that the complexity scheme could serve as a useful aid to inform the direction in which features of new items should be varied so as to reach a pre-determined or desired level of difficulty. The detailed scheme included in the appendix is a refinement of the original and continues to represent work in progress. Because of the recursive nature of the testing of this scheme (e.g., the same individuals wrote the scheme and rated the complexity of items), caution should be exercised in further interpretive use of the present version; further refinement and validation work is necessary.

External review. As part of work at this stage, the conceptual framework and some sample items were sent for review and comment by a panel of 16 experts from 9 countries. The reviews highlighted that there is a range of conceptions regarding the terrain covered by the term "numeracy", as expected in light of the conceptual analysis summarized in Section 1. Nonetheless, the reviews overall supported the conceptual framework for Numeracy assessment developed for the ALL project, and endorsed the approach to item production described earlier in this section.

7.2 Stage 2 (1999-2000): Additional item production, second feasibility study

Following the successful completion of Stage 1, additional items were contributed by experts from four countries: Austria, the Czech Republic, Hungary, and Sweden, in order to enrich the pool of Numeracy items and increase its cultural diversity. Some of these items were selected and adapted by the Numeracy team in order to fit the item production grid and the item development principles.

A small-scale feasibility study was conducted in the USA and The Netherlands, in which 44 additional items were tested on samples of about 55 cases in each country. Statistical analyses conducted at this stage were quite similar to those used during the first feasibility study, although the smaller sample prevented the use of IRT analysis.

Out of all the items tested in the two feasibility studies conducted in Stage 1 and Stage 2, a pool of 81 items was selected for further testing at the Pilot stage. These items had adequate psychometric characteristics, covered diverse levels of difficulty, encompassed key facets of the conceptual framework for Numeracy, and could be adapted without difficulty to a language other than English and to different units of measurement. These items did not show any remarkable error patterns indicating that stimuli, questions or task contexts were misunderstood. In addition, two easy Numeracy items were selected to become part of the “Core”, the screener test that is to be successfully passed by each respondent in order for him or her to receive one or more full-length test booklets with items assessing performance in the ALL skill domains.

7.3 Stage 3 (2001-2003): Preparations for Pilot testing in participating countries

Scoring materials. Once items for the Pilot stage were selected, the Numeracy team prepared detailed scoring guidelines for each item, as well as a training manual for scorers. In addition, the team prepared a detailed manual describing "critical elements" of each item that should be kept constant during the translation and adaptation work that came next.

Translation and adaptation for Pilot. The 81 items selected at the end of Stage 2 were translated and adapted by all participating countries, sometimes into multiple languages within the same country (e.g., Canada created English and French versions, and Switzerland created German, French, and Italian versions). The translation process was supported by training workshops held in 2000 and 2001 and by the support materials prepared by the Numeracy team, i.e., the manual describing critical elements of each item that should be kept constant across language versions, and a "translation and adaptation manual". Item adaptation aimed to maintain cognitive equivalency in terms of their task demands. Thus, for example, when units of measure or monetary values in some computational items were adapted to various country situations, guidelines emphasized the need to keep item demands comparable.

Further, each country prepared not only country-specific and language-specific sets of items, stimuli, and response pages, but also adapted the scoring instructions and the training manual for scorers. During this stage, all translators could also post questions by e-mail to a hotline staffed by members of the Numeracy team.

The next section provides an overview of the results from the Pilot phase and further explanations regarding the properties of the final 40 items selected for the Main assessment.

Part C: Pilot Study: Summary of Numeracy-related results and decisions

8. Overview and purposes of the Pilot stage

The ALL Pilot study was conducted in 2002 in six countries (and five languages): Belgium (French), Canada (French and English), Italy, Norway, Switzerland (German, French, Italian), and the United States. Large samples responded to test booklets assessing the key skill domains targeted by ALL: Document literacy, Prose literacy, Numeracy, and Problem-solving. The Background Questionnaire (BQ) provided additional information about respondents' experiences with Information Technology and about numerous variables of importance such as employment status, consumption of health services, participation in learning activities, and perceptions of quality of life. The BQ also collected additional personal information that could help to explain or interpret results pertaining to the key variables.

8.1 General goals

The Pilot study was designed to accomplish two goals:

a. Enable the ALL management and the development teams to evaluate the psychometric characteristics of the item pools created to assess the different skill domains, and provide sufficient data from which to select the items to be included in the test booklets and the BQ that will be used in the Main ALL survey.

b. Enable participating countries to field-test and evaluate all survey materials and administrative procedures planned for the Main ALL survey, such as contacting respondents, setting home visits, test administration, scoring of responses, data capture, and quality assurance mechanisms.

8.2 Numeracy-specific goals

In addition to the two general goals listed above, the Pilot study aimed to provide information regarding specific issues relevant to Numeracy assessment in ALL:

- Enable selection of 40 items out of the 81 Numeracy items tested in the Pilot, whose development was described in Part B. The target number of 40 items was determined on the basis of knowing average response times per item in the feasibility studies, and the overall time expected to be available for Numeracy assessment in the Main study.
- Analyze response patterns and verify that errors are not caused by unclear or problematic stimuli, questions, or instructions, or by inconsistent translation. Also, determine whether the multiple scoring codes created for some items should be maintained for the Main study
- Collect data on the extent to which respondents used the calculator provided as part of the assessment.
- Evaluate the relationships between Numeracy scale scores and variables measured by the BQ, including scores on scales related to numeracy practices and to beliefs and attitudes.

9. Methods

9.1 Respondents

In the ALL Pilot study, respondents aged 16 to 65+ were individually tested in their homes for approximately 90-120 minutes. In each participating country, around 1000-1500 individuals were tested overall, although each respondent was not tested in all of the domains in ALL. Given the goals of the Pilot study, respondents did not comprise a probability sample from each country, although care was taken to recruit individuals from diverse locations and stratify and balance the sample in terms of gender, age groups, and educational levels.

9.2 Procedure and instruments

All respondents were first given a short "Core", a screener test consisting of four simple Prose and Document Literacy tasks and two simple Numeracy tasks, which were read aloud by the examiner. Respondents were then given test booklets that included a subset of items from one or two of the key skill domains, as well as BQ questions. Respondents wrote their answers inside the test booklet. Those who received Numeracy items were free at all times to use a calculator and a ruler provided by the examiner.

Each respondent who received a Numeracy test answered only about half of the 81 items evaluated in the Pilot study. These items were divided into four blocks that were paired in various permutations into four booklets, thus each respondent received two of four blocks. This arrangement enabled testing of the complete Numeracy item pool and evaluation of the relationships between all items, although each respondent answered only half of the items. The total number that tried each Numeracy item averaged about 950 respondents across the countries and languages listed above.

9.3 Scoring

Items were scored by trained teams in each country according to the scoring rubrics and guidelines designed by the Numeracy team. Double-scoring procedures were implemented on a sample of test booklets in each country, to evaluate consistency of scoring and detect problems with scoring instructions. A third scorer arbitrated disagreements. During the scoring process, a member of the Numeracy team answered the questions from scorers in the participating countries on an active electronic listserve.

9.4 Data-analysis.

Various statistical analyses were conducted on the available data, with scaling based on Item Response Theory, analysis of difficulty levels of items and of scale characteristics using classical test theory, analysis of the frequency of different types of errors across countries, and correlational analysis of linkages between summary scores on Numeracy items and BQ variables and scales of interest. Key analyses were performed at the Educational Testing Service in Princeton, and some were conducted by members of the Numeracy team and by Statistics Canada.

10. Results

10.1 Items

With the ultimate goal of selecting 40 items for the Main Survey, the Numeracy team first looked at the Pilot data to identify items that may be problematic in terms of psychometric anomalies or scoring unreliability. Some of the key patterns or data examined were the following:

- Wide variation in performance on an item from country to country or across gender groups could mean that the context of the item was not universal or it could reveal discrepancies in either adaptation of items or test administration.
- When the observed performance on an item (for example, "percent correct" in classical test theory terms) deviates significantly from the difficulty level predicted from the theoretical complexity scheme (See Part A), it could indicate misunderstanding of the question or the presence of unexpected factors that cause response errors in some but not all countries.
- A large number of disagreements between scorers for an item could indicate that a scoring rubric was not discriminating properly. In addition, anecdotal reports from scorers on the listserv can be used to flag a few items whose scoring rubrics were difficult to score for various reasons.

Based on such considerations, the team rejected very few of the 81 items, as gross problems were mostly already eliminated after the two feasibility studies and translation and adaptation process. A few problems with specific items were revealed and addressed by technical recommendations. For example, instructions to use the ruler on certain items were made more explicit both for the respondent and the examiner. Recommendations were made on how to standardize the production of stimuli which required that respondents perform a measurement of length but which were not printed to the same exact measurements in different countries. In addition, the complex scoring rubrics developed and used up to the Pilot study for error analysis were collapsed into a simpler correct/incorrect classification. This approach was chosen in light of the advantages in terms of simplifying scoring processes and scorer decisions and after analyses showed no significant loss in the information provided about respondents' skills for the purposes of this survey.

10.2 Scale

The pilot results reaffirmed that the Numeracy items as a whole constitute a cohesive scale. Across all the countries, the mean Chronbach's Alpha coefficient, a measure of the internal consistency of a group of items intended to represent an underlying construct, averaged 0.88 across the four blocks of Numeracy items. A result above 0.80 generally indicates an underlying consistency in a scale designed to measure cognitive skills. An average Alpha of 0.88 is especially informative, considering that the Numeracy items encompass a range of facets of content, and vary in terms of their contexts, literacy demands and response requirements.

Additionally, the construct as defined was validated by the strong correlation between the predicted and observed difficulty of the items in the pilot. With a wider and more diverse population than the feasibility study, the pilot study empirical results (percent correct) were highly correlated ($r = -0.799$) with the theoretical predictions of difficulty determined with the complexity scheme (Figure 2 and Appendix 2). Again, the reader is advised that further validation and refinement of the detailed scheme is planned.

11. Selection of items for the Main Numeracy assessment

Following the above analyses, 40 items were chosen that together satisfy several requirements or goals that are outlined below. Information about the extent to which the selected 40 items satisfy these requirements is briefly described.

11.1 Facet coverage in scale

A primary goal for the items selected for the Main study is that they represent the various aspects of the four facets outlined as part of the numeracy construct, so that the scale overall offers respondents an opportunity to demonstrate a variety of numeracy skills, tested within a range of realistic contexts.

Mathematical Idea and Type of Response. The first focus in selecting items was on covering the two key facets of "mathematical idea" (Facet 3) and "type of response" (Facet 2) See Table 1, Part A. A desired mix of the categories within these two facets was created and is listed in Table 3, and items of varying levels of difficulty were chosen to fit within that mix.

It should be noted that the categories within the facets are not mutually exclusive; that is, an item whose primary task involves "dimension" may also require dealing with "quantity" and one that requires interpretation may also involve estimation. Hence, the percentages in Table 3 are not absolute, as items could be classified in more than one way, depending on which of their several key demands are taken into consideration.

Table 3: Expected distribution of items within facet categories

Content area	%	Type of Response	%
Quantity	30	Interpret	35
Dimension	30	Compute	30
Relations	10	Estimate	10
Data	20	Order	10
Change	10	Measure/ Count	10
		Model	5

Contexts. Another consideration in selecting items for the Main study was to achieve a desirable mix of the contexts from which the tasks are drawn (Facet 1 in Table 1). Nearly half, 45%, of the 40 items chosen are situated in Everyday Life contexts, 25% are from situations that represent participation in the Larger Society, 20% are from the world of Work, and 10% represent more formal tasks that may be helpful in Further Learning.

Representations. A final concern was to vary the types of stimuli that were included in the final scale. Each type of stimuli listed in the description (Facet 4 in Table 1) is present in the 40 items, i.e., pictures, numbers and symbols, formulae, diagrams and maps, graphs, tables, and, of course, text. With respect to text, the 40 items selected include stimuli (and questions) that span the range from minimum text dependency to higher levels of dependency, as shown in Table 4. The level of literacy/text dependence of the stimulus and the question is an important consideration, especially when comparisons are made to the QL scale used in IALS, where stimuli were often quite heavily text-based and required considerable literacy skills.

Table 4: Text aspects of items

Text dependency	Percentage of items
Low	37.5%
Medium	35.0%
High	27.5%

11.2 Difficulty Levels in Scale

A major goal of the ALL Main Survey is to profile the skill distribution of a country's population. The 42 items to be chosen to assess Numeracy in ALL (including the two Numeracy items in the Core screener) should represent a range of levels of difficulty, so that they can discriminate reliably between performance levels of respondents.

Table 5: Distribution of difficulty levels

Level	Number of items
1 (easy)	6
2	10
3	17
4	6
5 (difficult)	3

Tentative estimates of the performance level required of respondents to answer items in nationally representative samples in the Main study were derived on the basis of IRT analyses of the Pilot data and comparisons to data from the Prior IALS, and appear in Table 5. (While IALS and other large-scale comparative surveys of adults' skills represent item difficulty on a continuous scale, when reporting key results, they group items in terms of five difficulty levels, where level 1 refers to easy items and level 5 refers to difficult items). The 40 items (plus two Core items) selected for the Main appear to form a satisfactory distribution across five performance levels. The distribution of items in terms of difficulty levels provides the most information at the center of the expected population distribution, and thus promises a rich field of data from which to draw a profile of the population's numeracy.

11.3 Balance between blocks

Within the structure of the Main Survey, Numeracy items are organized by blocks. The BIB ("Balanced Incomplete Block") design adopted for the ALL survey, which was also used in IALS and other large-scale surveys, dictates that a block of Numeracy items be paired in booklets with a block of Prose and Document Literacy items or with a second block of Numeracy items. Average time requirements recorded during the pilot indicated that the allotment of 20 Numeracy items per block worked well and should be continued for the Main survey.

Hence, another requirement from the 40 items chosen for the Numeracy scale is that they can be divided into two blocks that reflect the difficulty level and the facet coverage of the entire scale, for psychometric as well as for "face validity" reasons. Subsequently, two 20-item blocks were created which are well balanced with respect to the means of the IRT indicators of discrimination (Mean Rbiserial) and difficulty (Mean b), and in terms of the number of items in each block that represent the five content areas listed in the conceptual framework of numeracy facets, as shown in Table 6 and Table 7, respectively. Thus, the two blocks that make up the Numeracy scale overall satisfy the requirement for psychometric equivalency and each of them represents the richness of the full construct.

Table 6: Blocks by psychometric indicators

	Block A	**Block B**
Discrimination (Mean Rbiserial)	0.63	0.67
Difficulty (Mean b)	-0.33	-0.33

Table 7: Blocks by content categories

	Block A	**Block B**
Quantity	6	6
Dimension	7	6
Relations	2	2
Data	3	4
Change	2	2

12. Summary

The results from the Pilot Study provide support both for the content validity and the construct validity of the Numeracy scale. The 40 items selected for the Main Study cover the Numeracy domain as envisioned. They reflect the facets contained in the definition of measurable numerate behavior, include items at all levels of difficulty, and satisfy psychometric requirements common in international comparative studies.

In addition, analysis of the pilot data suggest that "non-cognitive" items, including several item clusters developed by the Numeracy team, can serve as useful covariates of observed performance on Numeracy items and of numeracy skill levels. Some of these items ask about general attitudes concerning mathematical tasks and about recollections of mathematics learned in school settings. Other items ask about a respondent's involvement with particular numeracy tasks on the job, about self perceptions as to whether the respondent's math skills are sufficient to accomplish what is required at work, and examine certain beliefs, attitudes, and practices related to numeracy. Another item asks for a self-assessment of the extent to which a respondent used a calculator during the Numeracy block.

Overall, the forty items chosen for Numeracy assessment can be used to describe the distribution of numeracy skills in the population as well as in important sub-groups. In addition, the supplemental non-cognitive items included in the Background Questionnaire can shed light on possible antecedents and consequences of the numeracy skills that are of interest to policy makers and other audiences.

References

American College Testing. (1997). *Work keys: Preliminary technical handbook*. Iowa City, IA: Author.

Australian Education Council (The Mayer Committee). (1992). *The key competencies report: Putting general education to work.* Canberra: The Australian Education Council and Ministers for Vocational Education, Employment, and Training.

Baker, D., & Street, B. (1994). Literacy and numeracy: Concepts and definitions. In T. Husen & E. A. Postlethwaite (Eds.), *Encyclopedia of education*. New York: Pergamon Press.

Beazley, K. (1984). Education in Western Australia: Report of the Committee of Inquiry into Education in Western Australia. Education Department of Western Australia.

Benn, Roseanne (1997) *Adults count too: mathematics for empowerment*, Leicester, UK: National Institute of Adult Continuing Education (NIACE)

Bishop, A. (1988). Mathematical enculturation: a cultural perspective in mathematics education. Dordrecht, Holland: D. Reidel Publishing Co.

Buckingham, E. (1997). *Specific and generic numeracies of the workplace*. Melbourne: Deakin University.

Carnevale, A. P., Gainer, L. J., & Meltzer, A. S. (1990). *Workplace basics: The essential skills employers want*. San Francisco: Jossey-Bass.

Carraher, T. N., Schliemann, A. D., & Carraher, D. W. (1988). Mathematical concepts in everyday life. In G. B. Saxe & M. Gearhart (Eds.), *Children's Mathematics* (pp. 71-88). San Francisco: Jossey-Bass.

Coates et al. (1995). *National Reporting System*. Australia: Commonwealth of Australia and the Australian National Training Authority.

Coben, , D. (2000). Numeracy, mathematics, and adult learning. In I. Gal (Ed.), *Adult numeracy development: Theory, research, practice* (pp. 33-50). Cresskill, NJ: Hampton Press.

Cockcroft, W.H. (1982). Report of the Committee of Inquiry into the Teaching of Mathematics in Schools. London: HMSO.

Cumming, J., Gal, I., and Ginsburg, L. (1998). *Assessing mathematical knowledge of adult learners: Are we looking at what counts?* (Technical Report TR98-05). Philadelphia, PA: University of Pennsylvania, National Center on Adult Literacy.

Curry, D., Schmitt, M.J., and Waldron, S. (1996, July). A framework for adult numeracy standards: The mathematical skills and abilities adults need to be equipped for the future. Boston, MA: The Adult Numeracy Practitioners Network.

Dossey, J.A. (1997). "Defining and measuring quantitative literacy". In L.A. Steen (Ed.), *Why numbers count: Quantitative literacy for tomorrow's America*. New York: College Entrance Examination Board.

European Commission. (1996). *White paper on education and training: Teaching and learning towards the learning society*. Luxembourg: Office for Official Publications of the European Commission.

Fey, James T. (1990). "Quantity" In L.A. Steen (Ed.) *On the shoulders of giants: New approaches to numeracy*. Washington, DC: National Academy Press.

Forman, S. L. & Steen, L. A. (1999). *Beyond eight grade: Functional mathematics for life and work*. University of California, Berkeley: National Center for Research in Vocational Education.

Frankenstein, M. (1989). *Relearning mathematics: A different third 'R' – Radical maths*. London: Free Association Books.

Gal, I. (1993). *Issues and challenges in adult numeracy* (Technical Report TR93-15). Philadelphia, PA: National Center on Adult Literacy, University of Pennsylvania.

Gal, I. (1997). Numeracy: Imperatives of a forgotten goal. In L. A. Steen (Ed.), *Why numbers count: quantitative literacy for tomorrow's America* (pp. 36-44). New York: The College Board.

van Groenestijn, M. (1998). Constructive numeracy teaching as a gateway to independent learning. In Coben, D., and O'Donoghue, J. (Eds.), *Adults learning mathematics–4, Proceedings of the Fourth International Conference of Adults Learning Mathematics*. London: Goldsmiths University of London.

Heuvel-Panhuizen, M. van den, and Gravemeijer, K.P.E. (1991). Tests are not all that bad: An attempt to change the appearance of written tests in mathematics instruction at the primary school level. In L. Streefland (Ed.), *Realistic mathematics education in primary school*. The Netherlands: Utrecht.

Johnston, B. (1994, Summer). Critical numeracy? In *Fine print*, Vol.16, No. 4.

Jones, S. (1995). The distribution of literacy. In *Literacy, economy and society*. Ottawa: Statistics Canada.

Kirsch, I. S., Jungblut, A., & Mosenthal, P. B (1998). The measurement of adult Literacy. In S. T. Murray, I. S. Kirsch, & L. B. Jenkins (Eds.), *Adult literacy in OECD countries: Technical report on the first International Adult Literacy Survey* (pp. 105-134). Washington, DC: National Center for Education Statistics, U.S. Department of Education.

Kindler, Jan, et al. (1996). *Certificates in General Education for Adults*. Victoria, Melbourne: Adult, Community and Further Education Board.

Lave, J. (1998). Cognition in practice: Mind, mathematics, and culture in everyday life. Cambridge, UK: Cambridge University Press.

Leonelli, E., Merson, M.W., Schmitt, M.J., and Schwendeman, R. (Eds.). (1994). *The Massachusetts ABE math standards project* (Vols. 1 and 2). Holyoke, MA: Holyoke Community College.

Marr, B., and Tout, D. (1997). A numeracy curriculum: Australian Association of Mathematics Teachers (AAMT) conference proceedings. Melbourne: AAMT.

Mathematics Functional Expert Group-PISA Study, Jan de Lange, Chair. (1998, June). *Framework for assessing mathematical literacy*. Draft technical report submitted to the OECD/Program for International Student Assessment (PISA) project.

McLeod, D.B. (1992). Research on affect in mathematics education: A reconceptualization. In Grouws, D.A. (Ed.), Handbook of research on mathematics teaching and learning (pp. 575-596). New York: Macmillan.

Mullis, I. V.S., Martin, M. O., Beaton, A. E., Gonzalez, E. J., Kelly, D. L., & Smith, T. A. (1998). *Mathematics and science achievement in the final year of secondary school: IEA's Third International Mathematics and Science Study (TIMSS)*. Boston: Center for the Study of Testing, Evaluation, and Educational Policy.

Murnane, R., Willett, J., and Levy, F. (1995, March). *The growing importance of cognitive skills in wage determination* (Working Paper No. 50-76). Cambridge, MA: National Bureau of Economic Research.

National Council of Teachers of Mathematics (NCTM). (1989). *Curriculum and evaluation standards for school mathematics*. Reston, VA: Author.

National Research Council. (1989). *Everybody counts: A report to the nation on the future of mathematics education*. Washington, DC: National Academy Press.

Nunes, T. (1992). Ethnomathematics and everyday cognition. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 557-574). New York: Macmillan.

Oates, T. (1992). Developing and piloting the NCVQ core skill units: An outline of method and a summary of findings (Report 16). London: National Council on Vocational Qualifications (NCVQ).

Queensland Department of Education (1994). Unpublished manuscript

Resnick, L. B. (1987). Learning in school and out. *Educational Researcher*, *16,* 13-20.

Robitaille, D. F., and Travers, K. J. (1992). International studies of achievement in mathematics. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 687-709). New York: Macmillan.

Rogoff, B., & Lave, J. (Eds.). (1984). *Everyday cognition: Its development in social context*. Cambridge, MA: Harvard University Press.

Rutherford, F.J and Ahlgren, A. (1990). *Science for all Americans*. New York: Oxford University Press.

Saxe, G.B. (1992). Culture and cognitive development: Studies in mathematical understanding. Hillsdale, NJ: Lawrence Erlbaum Associates.

Saxe, G.B., Dawson, V., Fall, R., and Howard, S. (1996). Culture and children's mathematical thinking. In Sternberg, R, and Ben-Zeev, T. (Eds.), *The nature of mathematical thinking* (pp. 119-144). Hillsdale, NJ: Lawrence Erlbaum Associates.

Schliemann, A.D., and Acioly, N.M. (1989). Mathematical knowledge developed at work: The contribution of practice versus the contribution of schooling. *Cognition and Instruction*, 6, 185-221.

Secretary's Commission on Achieving Necessary Skills. (1991). *What work requires of schools: A SCANS Report for America 2000*. Washington, D.C.: U.S. Dept. of Labor.

Senechal, Majorie (1990) "Shape" In L.A. Steen (Ed.) *On the shoulders of giants: New approaches to numeracy*. Washington, DC: National Academy Press.

Statistics Canada and The Organisation for Economic Cooperation and Development. (1996). *Literacy, economy, and society: First results from the International Adult Literacy Survey.* Ottawa: Author.

Statistics Canada, and Organisation for Economic Cooperation and Development. (1997). *Literacy skills for the knowledge society: Further results from the International Adult Literacy Survey*. Ottawa: Author.

Steen, L. A. (Ed). (1990). *On the shoulders of giants: New approaches to numeracy*. Washington, D.C.: National Research Council.

Steen, L. A. (Ed). (1997). *Why numbers count: quantitative literacy for tomorrow's America*. New York: The College Board.

Stein, S.G. (1995). *Equipped for the future: A customer-driven vision for adult literacy and lifelong learning.* Washington, D.C.: National Institute for Literacy.

Tobias, S. (1993). *Overcoming math anxiety*. New York: Norton.

Willis, S. (1990). *Being numerate: What counts?* Victoria: Australian Council for Educational Research (ACER).

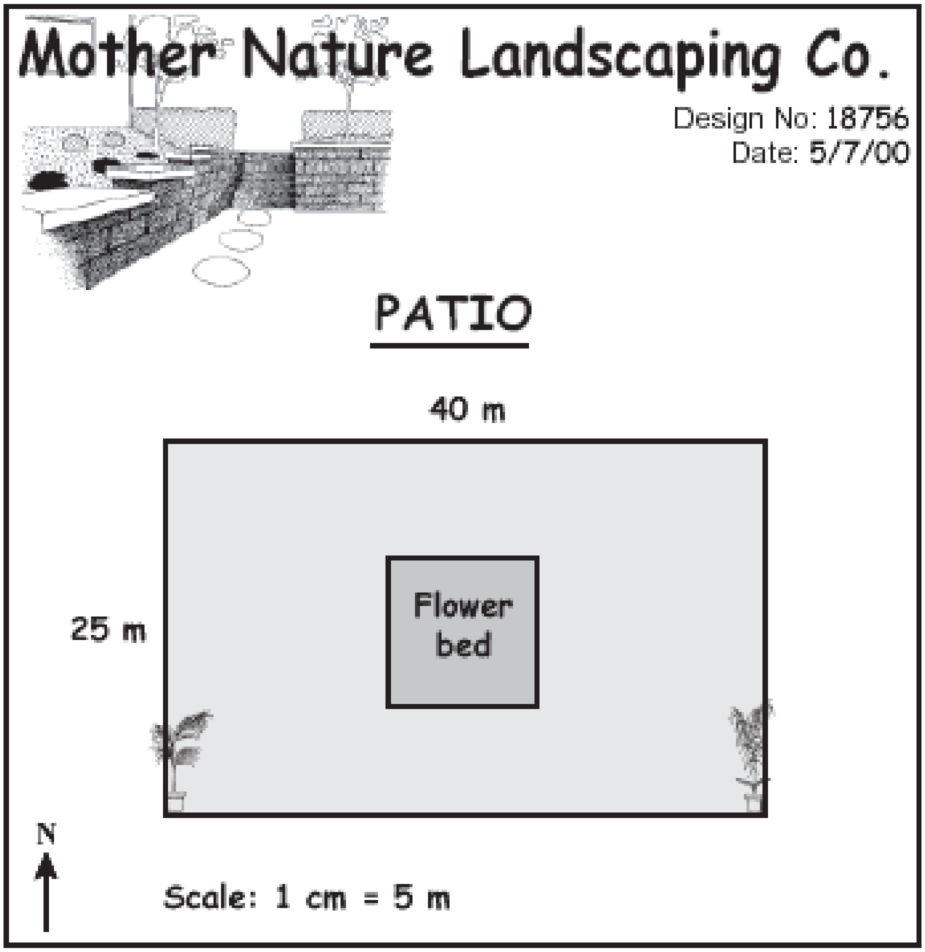
Appendix 1
Sample items

The sample items presented on the following pages show examples of the tasks that were developed for the pilot study. These are tasks that have been discarded for the main survey; however, they provide an idea of the types of tasks used in the final selection of items.

The tables accompanying them provide a broad analysis of each item that includes a few facet designations, the theoretical difficulty summary score that was assigned (ranging from 5, easiest, to 10, most difficult, and the difficulty statistics that were gathered during the pilot study.

Question
You wanted two pillows and bought them during the sale. In percent, how much did you save?

Item name	Content	Response	Other feature	Theoretical difficulty	% correct
Pillows	change	interpret	picture, low text	12	77%



Questions
Assume that you work for Mother Nature Landscaping Co.

You are given the sketch showing a patio that is to be landscaped by your company. The center square is to be planted with blooming flowers and the rest of the patio will contain a variety of potted plants arranged on its concrete surface.

a) The design calls for 5 large potted plants to be placed along the south edge of the patio, one in each corner (as shown) and 3 in between.
How far apart should the centers of the pots be if you want them to be equally spaced?

b) What is the area of the actual square flower bed in the middle of the patio?
(Note: you can use the scale on the bottom).

Item name	Content	Response	Other feature	Theoretical difficulty	% correct
Landscaping-a	dimension, space	interpret, compute	diagram	15	32%
Landscaping-b	dimension	measure, compute, use formula	diagram with scale	15	38%

Question.

Which pack of film gives you more for your money?

Explain how you decided.

Item name	Content	Response	Other feature	Theoretical difficulty	% correct
Best buy	quantity	compute/estimate explain	low text	7	66%

OPINION

LETTERS

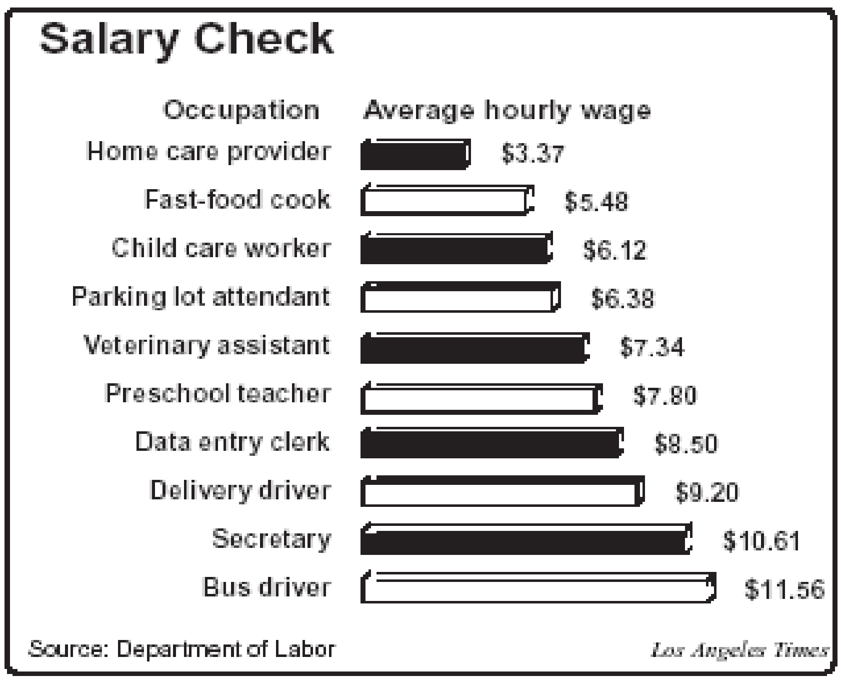
Last week we published the chart below on salaries. We received quite a few letters from readers. Some of the comments we include here.

From a secretary:
'I was interested to see your chart of average salaries. I was a homecare provider and changed jobs to a secretary last month.'

From an upset child care worker:
'The bus drivers have been on strike for 4 weeks for more pay. How dare they, the bus drivers already earn nearly double what child care workers earn per hour.'

And a final comment from a veterinary assistant who thinks we made an error:
'I am a veterinary assistant and make only $5.20 an hour. I think your chart must be wrong.'

Note from the editor: The source of data was the Department of Labor



Questions
Refer to the OPINION section of the newspaper.

a) What is the difference between the average hourly wage for a secretary and that of a home care provider?

b) How would you explain to the veterinary assistant that the average wage for veterinary assistants given in the chart can be correct?

Item name	Content	Response	Other feature	Theoretical difficulty	% correct
Wages-a	quantity, data	compute	locate in graph/text	10	78%
Wages-b	data	interpret, explain	Includes more text	15	27%

Appendix 2

Scoring for each of the Complexity Factors

Complexity Factor 1. Type of match/Problem transparency How difficult is it to identify and decide what action to take? How many literacy skills are required?		
Score 1	**Score 2**	**Score 3**
In the question and the stimulus, the information, activity or operation required: - is clearly apparent and explicit—and all required information is provided - is specified in little or no text, using familiar objects and/or photographs or other clear, simple visualizations - is about locating obvious information or relationships only - closed question—not open-ended	In the question and the stimulus, the information, activity or operation required: - is given using clear, simple sentences and/or visualizations where some translation or interpretation is required - is located within a number of sources within the text/activity. - fairly closed question	In the question and the stimulus, the information, activity or operation required: - is embedded in text where considerable translation or interpretation is required and/or - may need to be derived or estimated from a number of sources within or outside the text/activity and/or - the information or action required is not explicit or specified - more complex, open-ended task

Complexity Factor 2. Plausibility of distractors How many other pieces of mathematical information are present? Is all the necessary information there?		
Score 1	**Score 2**	**Score 3**
- no other mathematical information is present apart from that requested—no distractors	- there is some other mathematical information in the task that could be a distractor - the mathematical information given or requested can occur in more than one place - may need to bring to the problem simple information or knowledge from outside the problem.	- other irrelevant mathematical information appears - mathematical information given or requested appears in several places. - necessary information or knowledge is missing, so outside information or knowledge needs to be brought in

Complexity Factor 5. Expected number of operations How many steps and types of steps are required?		
score 1	**score 2**	**score 3**
one operation, action or process	application of two or three steps, the same or similar operation, action or process	integration of several steps covering more than one different operation, action or process

Complexity Factor 3. Complexity of mathematical information/answer required How complex is the mathematical information that needs to be manipulated?				
score 1	**score 2**	**score 3**	**score 4**	**score 5**
Context Based on very concrete, real life activities, familiar to most in daily life.	Based on common, real life activities.	Based on real life activities, but less often encountered.	Based on real life activities but unfamiliar to most	Based on abstract ideas or unfamiliar activity in a context new to most.
Quantity Whole numbers to 1,000 Fractions, decimals, percents - benchmark fractions ($^{1}/_{2}$, $^{1}/_{4}$, $^{3}/_{4}$) - decimal fraction for a half only (0.5) and equivalent as a percentage (50%)	- large whole numbers including millions - other benchmark fractions, like $^{1}/_{3}$ and $^{1}/_{10}$ - common decimals, like 0.1, 0.25 to 2 decimal places - common whole number percents, like 25% and 10%.	- large whole numbers including billions - other fractions - decimals to 3 decimal places (other than money) - all whole number percents	- negative integers - all remaining fractions, decimals and percentages	- all remaining types of rational (and some irrational) numbers including directed numbers
Pattern and relationship - very simple whole number relations and patterns	- simple whole number rates and ratios - whole number relations and patterns	- rates and ratios - relations and patterns including written everyday generalizations	- complex ratios, relations, patterns - simple formula	- formal mathematical information such as more complex formulae, knowledge of relationships between dimensions or variables, etc
Measures/ Dimension/Space - standard monetary values - common everyday measures for length (whole units) - time (dates, hours, minutes) - simple, common 2D shapes - simple localised maps or plans (no scales)	- everyday standard measures for length, weight, volume , including common fraction and decimal units - common 3D shapes and their representation via diagrams or photos - common types of maps or plans with visual scale indicators	- other everyday measures (area included) including fraction and decimal values - more complex 2D and 3D shapes, or a combination of 2 shapes - area and volume formulae - common types of maps or plans with ratio type scales	- all kinds of measurement scales - complex shapes or combinations of shapes	
Chance/Data - simple graphs, tables, charts with few parameters and whole number values - simple whole number data or statistical information in text	- graphs, tables, charts with common data including whole number percents—whole number scales in 1s, 2s, 5s or 10s - data or statistical information including whole number percents	- graphs, tables, charts with more complex data (not grouped data) - more complex data or statistical information including common average, chance and probability values - scales: more complex whole number, fractional or decimal	- complex graphs, tables or charts including grouped data - complex data or statistical information including probabilities, measures of central tendency and spread	

Complexity Factor 4. Complexity of Type of operation/skill How complex is the mathematical action that is required?				
score 1	**score 2**	**score 3**	**score 4**	**score 5**

Communicate no explanation - a single simple response required (orally, or in writing)	- no explanation - a simple response required (orally, or in writing	- simple explanation of a (level 1 or 2) mathematical process required (orally, or in writing) -	- explanation of a (level 3) mathematical process required (orally, or in writing)	- complex, abstract and generative reasoning or explanation required
Compute - a simple arithmetical operation (+, -, x, ÷) with whole numbers or money	- calculating common fraction, decimal fraction and percentages of values - using common rates (e.g. $/lb.); time calculations; etc - changing between common equivalent fraction, decimal and percent values, including for measurements e.g. $^1/_4$ kg = 0.250kg	- more complex applications of the normal arithmetical operations such as calculating with fractions and more complex rates, ratios, decimals, percentages, or variables - simple probability calculations	- applications of other mathematical operations such as squares, square roots, etc	- more advanced mathematical techniques and skills e.g. trigonometry
Estimate	- estimating and rounding off (when requested) to whole number values or monetary units	- estimating and rounding off to requested number of decimal places	- making a contextual judgment re whether a found answer is realistic or not and changing the answer to the appropriate correct rounded (but not necessarily mathematically correct) answer.	
Use formula/ model	- evaluating a given formula involving common operations (+, -, x, ÷)		- developing/creating and using straight forward formulae - using strategies such as working backwards or backtracking (e.g. 15% of ? = $255)	- generative reasoning - using and interpreting standard algebraic and graphical conventions and techniques
Measure - knowing common straight forward measures - naming, counting, comparing or sorting values or shapes	- visualizing and describing shapes, objects or geometric patterns or relationships - making and interpreting standard measurements using common measuring instruments	- using angle properties and symmetry to describe shapes or objects - estimating, making and interpreting measurements including interpolating values between gradations on scales - converting between standard measurement units within the same system	- calculating measures of central tendency and spread for non-grouped data - converting between non-standard measurement units within the same system - counting permutations or combinations	- converting between measurements across different systems
Interpret - locating/identifying data in texts, graphs and tables - orientating oneself to maps and directions such as right, left, etc	- reading and interpreting data from texts, graphs and tables - following or giving straight forward directions	- interpolating data on graphs - calculating distances from scales on maps	- generating, organising, graphing non-grouped data - extrapolating data - reading and interpreting trends and patterns in data on graphs, including slope/gradient	- graphing grouped data - calculating measures of central tendency and spread for grouped data