APPENDIX **A**

LITERACY PERFORMANCE ON THREE SCALES: DEFINITIONS AND RESULTS...............

IRWIN S. KIRSCH

Educational Testing Service,
Princeton, New Jersey, United States of America,
Telephone 1-609-734-1516 or e-mail ikirsch@ets.org

The methodology used to derive literacy skill levels was developed and tested for the International Adult Literacy Survey by Statistics Canada and the Educational Testing Service (ETS) in the United States of America. This appendix describes this methodology. Any requests for more detailed information should be directed to the author, Irwin Kirsch of the ETS.

The performance results for the International Adult Literacy Survey (IALS) are reported on three scales — prose, document and quantitative — rather than on a single scale. Each scale ranges from 0 to 500. Scale scores have, in turn, been grouped into five empirically determined literacy levels. Each of these levels implies an ability to cope with a particular subset of reading tasks. The balance of this appendix reports the proficiency achieved on each scale by adults in each participating country, and explains how to interpret this data by describing the scales and the kinds of tasks that were used in the test and the literacy levels that have been adopted.

While the literacy scales make it possible to compare the prose, document and quantitative skills of different populations and to study the relationships between literacy skills and various factors, the scale scores by themselves carry little or no meaning. In other words, whereas most people have a practical understanding of what it means when the temperature outside reaches 10°C, it is not intuitively clear what it means when a particular group is at 287 on the prose scale, or 250 on the document scale.

One way to gain some understanding about what it means to perform at various points along a literacy scale is to identify a set of variables that can be shown to underlie performance on these tasks. Collectively, these variables provide a framework for understanding what is being measured in a particular assessment and what skills and knowledge are being demonstrated by various levels of proficiency.

Toward this end, the appendix begins by describing how the literacy scale scores were defined. A detailed description of the prose, document and quantitative literacy scales is then provided, including a definition of each of the five levels and the percentages of adults in each of the participating countries demonstrating proficiency in each level. Sample tasks are described to illustrate the types of materials and task demands that characterise the five levels on each scale.

DEFINING THE LITERACY LEVELS

The item response theory (IRT) scaling procedures that were used in the IALS provide a statistical solution for establishing one or more scales for a set of tasks in which the ordering of difficulty is essentially the same for everyone. First, the difficulty of tasks is ranked on the scale according to how well respondents actually perform them. Next, individuals are assigned scores according to how well they do on a variety of tasks at different levels.

The scale point assigned to each task is the point at which individuals with that proficiency score have a given probability of responding correctly. In this survey, an 80% probability of correct response was the criterion used. This means that individuals estimated to have a particular scale score will consistently perform tasks — with an 80% probability — like those at that point on the scale. It also means they will have a greater than 80% chance of performing tasks that are lower than their estimated proficiency on the scale. It does not mean, however, that individuals with low proficiency can never succeed at more difficult tasks—that is, on tasks with difficulty values higher than their proficiencies. They may do so some of the time. Thus, it means that their probability of success is relatively low. In other words, the more difficult the task relative to their proficiency, the lower the likelihood of a correct response.

An analogy might help clarify this point. The relationship between task difficulty and individual proficiency is much like the high jump event in track and field, in which an athlete tries to jump over a bar that is placed at increasing heights. Each high jumper has a height at which he or she is proficient. That is, the jumper can clear the bar at that height with a high probability of success, and can clear the bar at lower heights almost every time. When the bar is higher than the athlete's level of proficiency, however, it is expected that the athlete will be unable to clear the bar consistently.

Once the literacy tasks are placed along each of the scales using the criterion of 80%, it is possible to see how well the interactions among various task characteristics explain the placement of tasks along the scales. Analyses of the interactions between the materials being read and the tasks based on these materials reveal that an ordered set of information-processing skills appears to be called into play to successfully perform the various tasks displayed along each scale (Kirsch and Mosenthal 1993).

To capture this order, each scale is divided into five levels reflecting the empirically determined progression of information-processing skills and strategies:

- Level 1 (0 to 225)
- Level 2 (226 to 275);
- Level 3 (276 to 325);
- Level 4 (326 to 375); and
- Level 5 (376 to 500).

It is worth noting that, while some of the tasks were at the low end of a scale and some at the very high end, most had values in the 200–400 range. It is also important to recognise that these levels were selected not as a result of any statistical property of the scales, but rather as the result of shifts in the skills and strategies required to succeed on various tasks along the scales, ranging from simple to complex.

The remainder of this appendix describes each scale in terms of the nature of task demands at each of the five levels. For each scale, sample tasks at each level are presented, and the factors contributing to their difficulty are discussed. The aim is to add meaning to the scales and to facilitate interpretation of the overall results as well as the breakdowns given in the main body of the report.

INTERPRETING THE PROSE LITERACY LEVELS

The ability to understand and use information contained in various kinds of textual material is an important aspect of literacy. The IALS therefore included an array of prose selections, including text from newspapers, magazines and brochures. The material varied in length, density, content, and use of structural or organisational aids such as headings, bullets and special typefaces. All prose samples were reprinted in their entirety with the original layout and typography intact.

Each prose selection was accompanied by one or more tasks or directives asking the reader to perform specific tasks. These tasks represent three major aspects of information-processing: locating, integrating and generating. Locating tasks require the reader to find information in the text based on conditions or features specified in the question or directive. The match may be literal or synonymous, or the reader may need to make an inference in order to perform successfully. Integrating tasks ask the reader to pull together two or more pieces of information in the text. In some cases the information can be found in a single paragraph, while in others it appears in different paragraphs or sections. In the generating tasks, readers must produce a written response by processing information from the text and also by making text-based inferences or drawing on their own background knowledge.

In all, the prose literacy scale includes 34 tasks with difficulty values ranging from 188 to 377. These tasks are distributed by level as follows: Level 1 (5 tasks); Level 2 (9 tasks); Level 3 (14 tasks); Level 4 (5 tasks); and Level 5 (1 task). It is important to remember that the tasks requiring the reader to locate, integrate and generate information extend over a range of difficulty as a result of interactions with other variables including:

- the number of categories or features of information the reader must process

- the extent to which information given in the question or directive is obviously related to the information contained in the text

- the amount and location of information in the text that shares some of the features with the information being requested and thus, seems plausible but does not fully answer the question; these are called 'distractors'

- the length and density of the text.

The five levels of prose literacy are defined on the following pages.

Prose Level 1 — Score range: 0 to 225

Tasks at this level require the reader to locate and match a single piece of information in the text. Typically the match between the task and the text is literal, although sometimes a low-level inference may be necessary. The text is usually brief or has organisational aids such as paragraph headings or italics that suggest where in the text the reader should search for the specified information. Generally, the target word or phrase appears only once in the text.

The easiest task in Level 1 (difficulty value of 188) directs respondents to look at a medicine label to determine the 'maximum number of days you should take this medicine.' The label contains only one reference to number of days and this information is located under the heading 'DOSAGE.' The reader must go to this part of the label and locate the phrase 'not longer than 7 days.'

MEDCO ASPIRIN   500

INDICATIONS: Headaches, muscle pains, rheumatic pains, toothaches, earaches. RELIEVES COMMON COLD SYMPTOMS.

DOSAGE: ORAL. 1 or 2 tablets every 6 hours, preferably accompanied by food, for not longer than 7 days. Store in a cool, dry place.

CAUTION: Do not use for gastritis or peptic ulcer. Do not use if taking anticoagulant drugs. Do not use for serious liver illness or bronchial asthma. If taken in large doses and for an extended period, may cause harm to kidneys. Before using this medication for chicken pox or influenza in children, consult with a doctor about Reyes Syndrome, a rare but serious illness. During lactation and pregnancy, consult with a doctor before using this product, especially in the last trimester of pregnancy. If symptoms persist, or in case of an accidental overdose, consult a doctor. Keep out of reach of children.

INGREDIENTS: Each tablet contains 500 mg acetylsalicylic acid.
Excipient c.b.p.   1 tablet.
Reg. No. 88246

Made in Canada by STERLING PRODUCTS, INC.
1600 Industrial Blvd., Montreal, Quebec H9J 3P1

0   6 7736 11079

Reprinted by permission

Prose Level 2 — Score range: 226 to 275

Like the tasks at Level 1, most of the tasks at Level 2 ask the reader to locate information. However, more varied demands are placed on the reader in terms of the number of responses the question requires, or in terms of the distracting information that may be present. For example, a task based on an article about the impatiens plant asks the reader to determine what happens when the plant is exposed to temperatures of 14°C or lower. A sentence under the section 'General care' states that 'When the plant is exposed to temperatures of 12–14°C, it loses its leaves and won't bloom anymore.' This task received a difficulty value of 230, just in the Level 2 range. What made this task somewhat harder than those identified at Level 1 is that the previous sentence in the text contains information about the requirements of the impatiens plant in various temperatures. This information could have distracted some readers, making the task slightly more difficult.

A similar task involving the same text asks the reader to identify 'what the smooth leaf and stem suggest about the plant.' The second paragraph of the article is labelled 'Appearance' and contains a sentence that states, '. . . stems are branched and very juicy, which means, because of the tropical origin, that the plant is sensitive to cold.' This

sentence distracted some readers from the last sentence in the paragraph: 'The smooth leaf surfaces and the stems indicate a great need of water.' This task received a difficulty value of 254, placing it in the middle of Level 2.

IMPATIENS

Like many other cultured plants, impatiens plants have a long history behind them. One of the older varieties was sure to be found on grandmother's windowsill. Nowadays, the hybrids are used in many ways in the house and garden.

Origin: The ancestors of the impatiens, *Impatiens sultani* and *Impatiens holstii*, are probably still to be found in the mountain forests of tropical East Africa and on the islands off the coast, mainly Zanzibar. The cultivated European plant received the name *Impatiens walleriana*.

Appearance: It is a herbaceous bushy plant with a height of 30 to 40 cm. The thick, fleshy stems are branched and very juicy, which means, because of the tropical origin, that the plant is sensitive to cold. The light green or white speckled leaves are pointed, elliptical, and slightly indented on the edges. The smooth leaf surfaces and the stems indicate a great need of water.

Bloom: The flowers, which come in all shades of red, appear plentifully all year long, except for the coldest months. They grow from "suckers" (in the stem's "armpit").

Assortment: Some are compact and low-growing types, about 20 to 25 cm. high, suitable for growing in pots. A variety of hybrids can be grown in pots, window boxes, or flower beds. Older varieties with taller stems add dramatic colour to flower beds.

General care: In summer, a place in the shade without direct sunlight is best; in autumn and spring, half-shade is best. When placed in a bright spot during winter, the plant requires temperatures of at least 20°C; in a darker spot, a temperature of 15°C will do. When the plant is exposed to temperatures of 12-14°C, it loses its leaves and won't bloom anymore. In wet ground, the stems will rot.

Watering: The warmer and lighter the plant's location, the more water it needs. Always use water without a lot of minerals. It is not known for sure whether or not the plant needs humid air. In any case, do not spray water directly onto the leaves, which causes stains.

Feeding: Feed weekly during the growing period from September to March.

Repotting: If necessary, repot in the spring or in the summer in light soil with humus (prepacked potting soil). It is better to throw the old plants away and start cultivating new ones.

Propagating: Slip or use seeds. Seeds will germinate in ten days.

Diseases: In summer, too much sun makes the plant woody. If the air is too dry, small white flies or aphids may appear.


Prose Level 3 — Score range: 276 to 325

Tasks at Level 3 on the prose scale tend to require the reader to search for information that requires low-level inferences or that meets conditions stated in the question. Sometimes the reader needs to identify several pieces of information that are located in different sentences or paragraphs rather than in a single sentence. Readers may also be asked to integrate or to compare and contrast information across paragraphs or sections of text.

A task at this level (with a difficulty value of 281) refers the reader to a page from a bicycle owner's manual to determine how to check to make sure the seat is in the proper position. The reader must locate the section labelled 'Fitting the Bicycle.' Then readers must identify and summarise the correct information in writing, making sure the conditions stated are contained in their summary.

A second Level 3 task, receiving a difficulty value of 310, directs the reader to look at a set of four movie reviews to determine which review was least favourable. Unlike some reviews that rate movies by points or some graphic such as stars, these reviews contain no such indicators. The reader needs to glance at the text of each review to compare what the reviewer said in order to judge which movie received the worst rating.

PROPER FRAME FIT

RIDER MUST BE ABLE TO STRADDLE BICYCLE WITH AT LEAST 2 cm CLEARANCE ABOVE THE HORIZONTAL BAR WHEN STANDING.

NOT LESS THAN 2cm

NOT LESS THAN 2cm
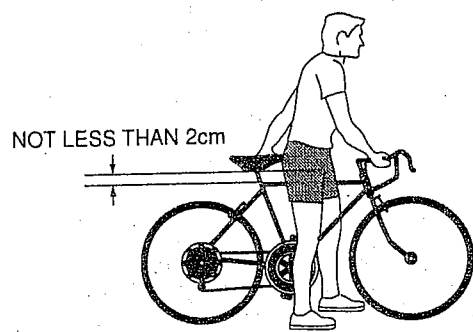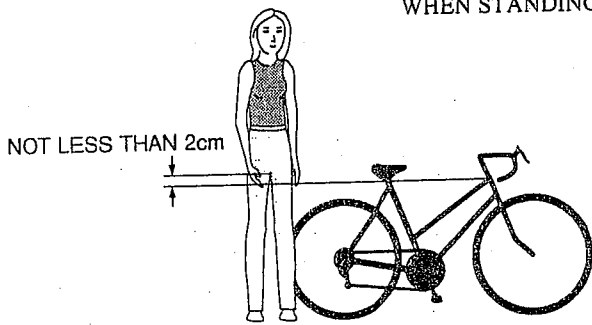
NOTE: Measurement for a female should be determined using a men's model as a basis.

PROPER SIZE OF BICYCLE	
FRAME SIZE	LEG LENGTH OF RIDER
430mm	660mm-760mm
460mm	690mm-790mm
480mm	710mm-790mm
530mm	760mm-840mm
560mm	790mm-860mm
580mm	810mm-890mm
635mm	860mm-940mm

OWNER'S RESPONSIBILITY

1. **Bicycle Selection and Purchase:** Make sure this bicycle fits the intended rider. Bicycles come in a variety of sizes. Personal adjustment of seat and handlebars is necessary to assure maximum safety and comfort. Bicycles come with a wide variety of equipment and accessories . . . make sure the rider can operate them.

2. **Assembly:** Carefully follow all assembly instructions. Make sure that all nuts, bolts and screws are securely tightened.

3. **Fitting the Bicycle:** To ride safely and comfortably, the bicycle must fit the rider. Check the seat position, adjusting it up or down so that with the sole of rider's foot on the pedal in its lowest position the rider's knee is slightly bent.

Note: Specific charts illustrated at left detail the proper method of determining the correct frame size.

The manufacturer is not responsible for failure, injury, or damage caused by improper completion of assembly or improper maintenance after shipment.

The Job Interview

Pre-interview

Try to learn more about the business. What products does it manufacture or services does it provide? What methods or procedures does it use? This information can be found in trade directories, chamber of commerce or industrial directories, or at your local employment office.

Find out more about the position. Would you replace someone or is the position newly created? In which departments or areas would you work? Collective agreements describing various standardised positions and duties are available at most local employment offices. You can also contact the appropriate trade union.

The Interview

Ask questions about the position and the business. Answer clearly and accurately all questions put to you. Bring along a note pad as well as your work and training documents.

The Most Common Types of Interview

One-on-one: Self explanatory.
Panel: A number of people ask you questions and then compare notes on your application.
Group: After hearing a presentation with other applicants on the position and duties, you take part in a group discussion.

Post-interview

Note the key points discussed. Compare questions that caused you difficulty with those that allowed you to highlight your strong points. Such a review will help you prepare for future interviews. If you wish, you can talk about it with the placement officer or career counsellor at your local employment office.

Another Level 3 question involved an article about cotton nappies. Here readers were asked to write three reasons why the author prefers to use cotton nappies over disposable nappies. This task was relatively difficult (318) because of several variables. First, the reader has to provide several answers requiring text-based inferences. Nowhere in the text does the author say, 'I prefer cotton nappies because ... '. These inferences are made somewhat more difficult because the type of information being requested is a 'reason' rather than something more concrete such as a date or person. And finally, the text contains information that may distract the reader

Prose Level 4 — Score range: 326 to 375

These tasks require readers to perform multiple-feature matching or to provide several responses where the requested information must be identified through text-based inferences. Tasks at this level may also require the reader to integrate or contrast pieces of information, sometimes presented in relatively lengthy texts. Typically, these texts contain more distracting information and the information that is requested is more abstract.

One task falling in the middle of Level 4 with a difficulty value of 338 directs readers to use the information from a pamphlet about a hiring interview to 'write in your own words one difference between the panel interview and the group interview.' Here readers needed to read the brief descriptions about each type of interview. And, rather than merely locating a fact about each or identifying a similarity, they need to integrate what was being presented to infer a characteristic on which the two types of interviews differ. Experience from other large-scale assessments reveals that tasks in which readers are asked to contrast information are more difficult, on average, than tasks in which they are asked to compare information to find similarities.

Prose Level 5 — Score range: 376 to 500

Typically tasks at this level require the reader to search for information in dense text that contains a number of plausible distractors. Some require readers to make high-level inferences or use specialised knowledge.

One task used in this assessment fell in Level 5. This task, receiving a difficulty value of 377, requires the reader to look at an announcement from a personnel department and 'list two ways in which CIEM helps people who will lose their jobs because of a departmental reorganisation.' The correct response requires readers to search through this text to locate the embedded sentence 'CIEM acts as a mediator for employees who are threatened with dismissal resulting from reorganisation, and assists with finding new positions when necessary.' This task is difficult because the announcement is organised around information that is different from what is being requested in the question. Thus, while the correct information is located in a single sentence, this information is embedded under a list of headings describing CIEM's activities for employees looking for other work. This list of headings serves as an excellent set of distractors for the reader who does not search for or locate the phrase containing the conditional information stated in the directive; that is, those who lose their jobs because of a departmental reorganisation.

AUSCO

AUSCO Manufacturing Company
Personnel Department

Centre on Internal and External Mobility

What is CIEM?

CIEM stands for Centre on Internal and External Mobility, an initiative of the personnel department. A number of workers of this department work in CIEM, together with members from other departments and outside career consultants.

CIEM is available to help employees in their search for another job inside or outside the Ausco Manufacturing Company.

What does CIEM do?

CIEM supports employees who are seriously considering other work through the following activities:
• **Job Data Bank**
After an interview with the employee, information is entered into a data bank that tracks job seekers and job openings at Ausco and at other manufacturing companies.
• **Guidance**
The employee's potential is explored through career counselling discussions.
• **Courses**
Courses are being organised (in collaboration with the department for information and training) that will deal with job search and career planning.
• **Career Change Projects**
CIEM supports and coordinates projects to help employees prepare for new careers and new perspectives.
• **Mediation**
CIEM acts as a mediator for employees who are threatened with dismissal resulting from reorganisation, and assists with finding new positions when necessary.

How much does CIEM cost?

Payment is determined in consultation with the department where you work. A number of services of CIEM are free. You may also be asked to pay, either in money or in time.

How does CIEM work?

CIEM assists employees who are seriously considering another job within or outside the company.

That process begins by submitting an application. A discussion with a personnel counsellor can also be useful. It is obvious that you should talk with the counsellor first about your wishes and the internal possibilities regarding your career. The counsellor is familiar with your abilities and with developments within your unit.

Contact with CIEM in any case is made via the personnel counsellor. He or she handles the application for you, after which you are invited to a discussion with a CIEM representative.

For more information

The personnel department can give you more information.

INTERPRETING THE DOCUMENT LITERACY LEVELS

Adults often encounter materials such as tables, schedules, charts, graphs, maps and forms at home, at work, or when travelling in their communities. The knowledge and skills needed to process information contained in these documents is therefore an important aspect of being literate in a modern society. Success in processing documents appears to depend at least in part on the ability to locate information in a variety of displays, and to use this information in various ways. Sometimes procedural knowledge may be required to transfer information from one source to another, as is necessary in completing applications or order forms.

The IALS document literacy scale contains 34 tasks that are ordered along the scale from 182 to 408 as the result of responses of adults from each of the participating countries. These tasks are distributed as follows: Level 1 (6 tasks); Level 2 (12 tasks); Level 3 (13 tasks); Level 4 (2 tasks); and Level 5 (1 task). By examining tasks associated with these proficiency levels, characteristics that are likely to make particular document tasks more or less difficult can be identified. Questions or directives associated with the various document tasks are basically of four types: locating, cycling, integrating and generating. Locating tasks require the reader to match one or more features of information stated in the question to either identical or synonymous information given in the document. Cycling tasks require the reader to locate and match one or more features of information, but differ from locating tasks because they require the reader to engage in a series of feature matches to satisfy conditions given in the question. The integrating tasks typically require the reader to compare and contrast information in adjacent parts of the document. In the generating tasks, readers must produce a written response by processing information found in the document and by making text-based inferences or drawing on their own background knowledge.

As with the prose tasks, each type of question or directive associated with a document task extends over a range of difficulty as a result of interactions among several other characteristics:

- the number of categories or features of information in the question the reader must process or match;

- the number of categories or features of information in the document that seem plausible or correct because they share some but not all of the information with the correct answer;

- the extent to which the information asked for in the question is obviously related to the information stated in the document; and

- the structure and content of the document.

A more detailed discussion of the five levels of document literacy follows.

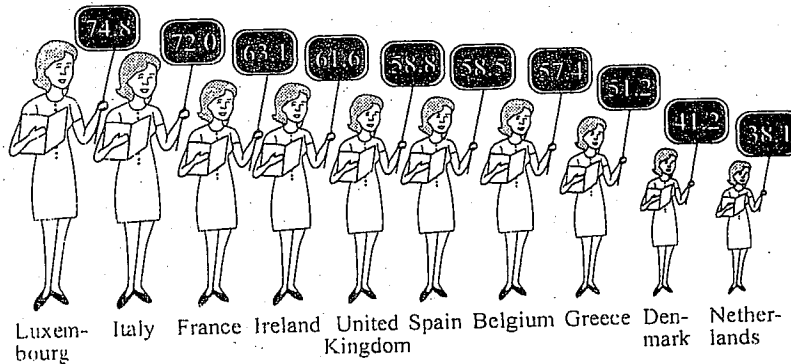Document Level 1 · Score range: 0 to 225

Most of the tasks at this level require the reader to locate a piece of information based on a literal match. Distracting information, if present, is typically located away from the correct answer. Some tasks may direct the reader to enter personal information onto a form.

One document task meeting this description (187) directs the reader to identify from a chart the percentage of teachers from Greece who are women. The chart displays the percentages of women teachers from various countries. Only one number appears on the chart for each country.

A very similar task involves a chart displayed in a newspaper showing the expected amounts of radioactive waste by country. This task, which has a difficulty value of 218, directs the reader to identify the country that is projected to have the smallest amount of waste by the year 2000. Again, there is only one percentage associated with each country. In this task, however, the reader must first identify the percentage associated with the smallest amount of waste and then match it to the country.

FEW DUTCH WOMEN AT THE BLACKBOARD

There is a low percentage of women teachers in the Netherlands compared to other European countries. In most of the other countries, the majority of teachers are women. However, if we include the figures for inspectors and school principals, the proportion shrinks considerably and women are in a minority everywhere.

| Luxem-bourg	Italy	France	Ireland	United Kingdom	Spain	Belgium	Greece	Den-mark	Nether-lands
74.8	72.0	63.1	61.6	58.8	58.5	57.1	51.2	41.2	38.1

Percentage of women teachers (kindergarten, primary, and secondary).

Document Level 2 — Score range: 226 to 275

Document tasks at this level are a bit more varied. While some still require the reader to match on a single feature, more distracting information may be present or the match may require a low-level inference.

One Level 2 task on the document scale (242) seems very similar to one described above for Level 1. This task directs the reader to use a chart to identify the year in which the fewest people in the Netherlands were injured by fireworks. Part of what may have made this task somewhat more difficult is that two charts were presented instead of just one. One, labelled 'Fireworks in the Netherlands,' depicts years and numbers representing funds spent in millions of U.S. dollars, while the other, 'Victims of fireworks,' uses a line to show numbers of people treated in hospitals. It is worth noting that in the second assessment, this label was changed to read 'number injured.'

Fireworks in the Netherlands

in millions of Australian dollars

Victims of fireworks

number injured
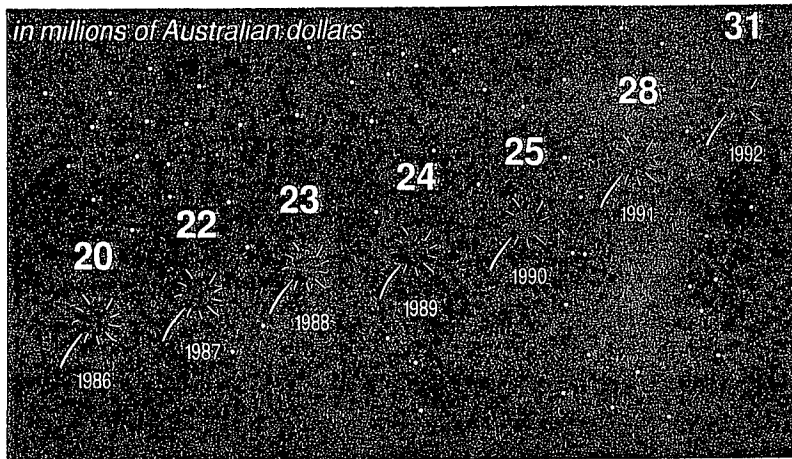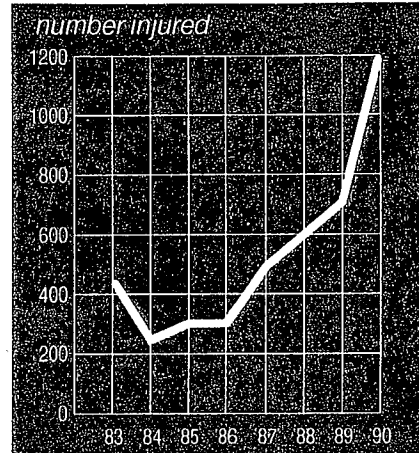
Several other tasks falling within Level 2 direct the reader to use information given to complete a form. In one case they are asked to fill out an order form to purchase tickets to see a play on a particular day, at a particular time. In another, readers are asked to complete the availability section of an employment application based on information provided that included: total number of hours they are willing to work, how they heard about the job, and availability of transportation.

Document Level 3 — Score range: 276 to 325

Tasks at this level appear to be most varied. Some require the reader to make literal or synonymous matches, but usually the matches require the reader to take conditional information into account or to match on multiple features of information. Some tasks at this level require the reader to integrate information from one or more displays of information. Other tasks ask the reader to cycle through a document to provide multiple responses.
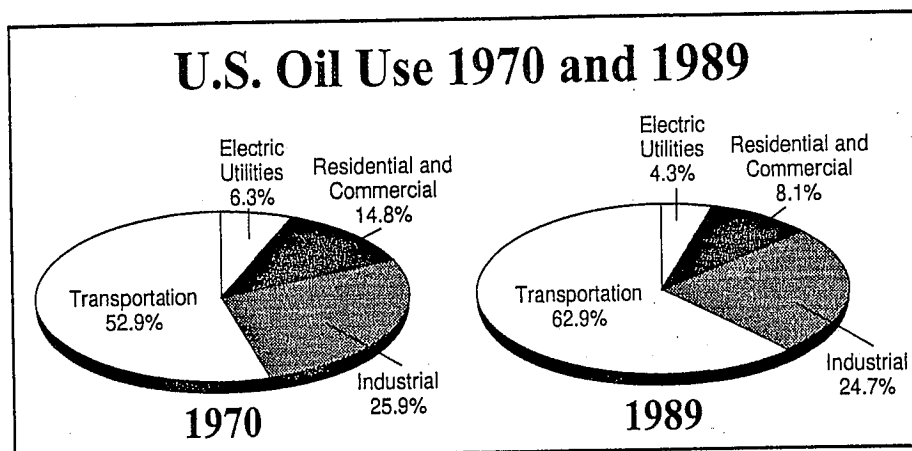
One task falling around the middle of Level 3 in difficulty involves the fireworks charts discussed above. This task directs the reader to write a brief description of the relationship between sales and injuries based on the information shown in the two graphs. This task received a difficulty value of 295. A second task, falling at high end of Level 3 (321), involves the use of a quick copy printing requisition form that might be found in the workplace. The task asks the reader to explain whether or not the quick copy centre would make 300 copies of a statement that is 105 pages long. In responding to this directive, the reader must determine whether conditions stated in the question meet those provided in the guidelines to this document.

QUICK COPY Printing Requisition

FILL IN ALL INFORMATION REQUESTED

GUIDELINES: This requisition may be used to order materials to be printed BLACK INK only, and in the quantities that are listed at the right.

■ SINGLE SHEET PRINTED 1 OR 2 SIDES — 2000 copies maximum
■ MORE THAN ONE SHEET UP TO 100 PAGES — 400 copies maximum
OVER 100 PAGES — 200 copies maximum

1. PROJECT TO BE CHARGED

2. TODAY'S DATE

3. TITLE OR DESCRIPTION

4. DATE DELIVERY REQUIRED

5. DO NOT MARK IN SHADED BOXES

| | X | | = | |
| NUMBER OF ORIGINALS | | NUMBER OF COPIES TO BE PRINTED | | TOTAL NUMBER OF IMPRESSIONS |

6. NUMBER OF SIDES TO BE PRINTED (Tick one box.)   1 ☐ One side   2 ☐ BOTH sides

7. COLOR OF PAPER (Fill in only if NOT white.) _____

8. SIZE OF PAPER (Fill in only if NOT 8½ x 11) _____

9. Tick any that apply:
   ☐ COLLATE

   BINDING: ☐ One staple at upper left
   ☐ Two staples in left margin
   ☐ BIND-FAST: ☐ Black
   ☐ Brown
   ☐ 3-hole punch

☐ Other instructions _____

AUTHORISATION AND DELIVERY

10. Project Director (print name) _____

11. Requisitioner (print your own name and phone no.) _____   extension

12. Tick one:

   ☐ Requisitioner will PICK UP completed job.

   ☐ Mail completed job to: _____   ROOM NO.
   Print name and room number

13. **KEEP PINK COPY at least 3 months.** When requesting information, you must refer to the requisition number printed here.

140468

QUICK COPY REGISTRATION NUMBER

D1320-03116 · 000000 · 000000

Document Level 4 — Score range: 326 to 375

Tasks at this level, like those in the previous levels, ask the reader to match on multiple features of information, to cycle through documents, and to integrate information; frequently however, these tasks require the reader to make higher order inferences to arrive at the correct answer. Sometimes, conditional information is present in the document, which must be taken into account by the reader.

One of the two tasks falling at this level (341) asks the reader to look at two pie charts showing oil use for 1970 and 1989. The question directs the reader to summarise how the percentages of oil used for different purposes changed over the period specified. Here the reader must cycle through the two charts, comparing and contrasting the percentages for each of the four stated purposes. Then the reader must generate a statement that captures these changes.

U.S. Oil Use 1970 and 1989

1970

Electric Utilities 6.3%
Residential and Commercial 14.8%
Transportation 52.9%
Industrial 25.9%

1989

Electric Utilities 4.3%
Residential and Commercial 8.1%
Transportation 62.9%
Industrial 24.7%

Document Level 5 — Score range: 376 to 500

Tasks at this level require the reader to search through complex displays of information that contain multiple distractors, to make high-level inferences, process conditional information, or use specialised knowledge.

The only Level 5 task in this international assessment involved a page taken from a consumer magazine rating clock radios. The most difficult task (408) involving this document asked the reader for the average advertised price for the basic clock radio receiving the highest overall score. This task required readers to process two types of conditional information. First, they needed to identify the radio receiving the highest overall score while distinguishing among the three types of clock radios reviewed: full-featured, basic and those with a cassette player. Second, they needed to locate a price. In making this final match, they needed to notice that two prices were given; the first, the suggested retail and the second, the average advertised price.

A second and considerably easier task involving this document and falling at the low end of Level 4 (327) asks the reader 'which full-featured radio is rated the highest on performance.' Again, readers needed to find the correct category of clock radio. Yet, they needed to process fewer conditions. Here they only needed to distinguish between the rating for 'Overall Score' and 'Performance.' It is possible that some adults identified the full-featured radio as receiving the highest 'Overall Score' rather than the one rated highest in 'Performance' as specified in the question. As such, 'Overall Score' would be considered a plausible distractor. Another factor that likely contributed to this task's difficulty is that 'Overall Score' is given a numerical value while the other features are rated by a symbol. It may be that some adults found the correct category ('Performance'), but selected the first radio listed, assuming it performed best. The text accompanying this table indicates the radios are rated within a category by overall score. It is easy to imagine that some people may have equated overall score with overall performance.

RATINGS

Better ◄——————————► Worse
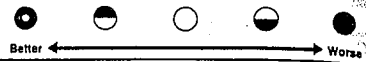
Clock radios

Listed by types; within types, listed in order of overall score. Differences in score of 4 points or less were not deemed significant.

1 Brand and model. If you can't find a model, call the company. Phone numbers are listed on page 736.

2 Price. The manufacturer's suggested or approximate retail price, followed by the average advertised price.

3 Dimensions. To the nearest centimetre.

4 Overall score. A composite, encompassing all our tests and judgments. A "perfect" radio would have earned 100 points.

5 Convenience. This composite judgment reflects such things as the legibility of the display, the ease of tuning the radio and setting the alarm, and the presence or absence of useful features.

6 Performance. An overall judgment reflecting performance in our tests of: sensitivity and selectivity; tuning ease; capture ratio, the ability to bring in the stronger of two stations on the same frequency; image rejection, the ability to ignore signals from just above the band, resistance to interference from signals bouncing off aircraft and such.

7 Sensitivity. How well each radio received a station with little interference.

8 Selectivity. How well each radio received clearly a weak station next to a strong one on the dial.

9 Tone quality. Based mainly on computer analysis of the speaker's output and on listening tests, using music from CDs. No model produced high-fidelity sound.

10 Reversible time-setting. This useful feature makes setting clock and alarm times easy. If you overshoot the desired setting, you simply back up.

11 Dual alarm. Lets you set two separate wake-up times.

1 Brand and model	2 Price	3 Dimensions HxWxD, cm.	4 Overall Score	5 Convenience	6 Performance	7 Sensitivity	8 Selectivity	9 Tone quality	10 Reversible time setting	11 Dual alarm	Warranty, months	Advantages	Disadvantages	Comments
Full-featured clock radios														
RCA RP-3690	$50/$40	8x25x18	86	◓	◑	◑	◑	◓	✔	✔	12	A,B,D,H,J,L,O,T,U		A
Sony ICF-C303	50/45	5x20x15	84	◓	◑	◓	○	◒	✔	✔	12	C,E,F,I,N,T		C
Panasonic RC-X220	50/45	10x28x13	82	◑	◑	◑	◑	○	✔	✔	12	A,G,K,M,O,S,T,U	b,c	A
Realistic 272	50/30	5x28x15	79	◑	○	◑	◑	◒	✔	✔	3	A,G,H,K,O,T		D
Magnavox AJ3900	65/—	15x38x13	78	○	◓	◑	◓	◑	—	✔	3	D,G,K,M,O,R,T	b,g	B
Emerson AK2745	39/20	8x28x15	70	○	◒	◑	◒	○	✔	✔	3	G,O	g	K
Soundesign 3753	20/20	8x23x13	62	○	◒	●	○	○	✔	✔	3	J,Q	d,h	J
Basic clock radios														
Realistic 263	28/18	10x20x10	74	○	◑	◑	○	◒	—	—	3	A,D,H,O,P,U	h	—
Soundesign 3622	12/10	5x20x13	68	◒	◑	◓	◒	◒	—	—	3	U	d	L
Panasonic RC-6064	18/15	5x20x13	67	◒	◑	◓	○	○	—	—	12	—	b,c	—
General Electric 7-4612	13/10	5x20x13	66	◒	○	◓	◒	○	—	—	12	A,D	a,g	—
Lloyds CR001	20/15	5x18x13	64	◒	○	◓	○	◒	—	—	3	U	—	—
Sony ICF-C240	15/13	5x18x15	63	◒	○	○	○	○	—	—	12	—	f,g	—
Emerson AK2720	19/10	5x20x13	61	◒	○	◑	●	○	—	—	3	O,T	e	K
Gran Prix D507	15/10	5x18x10	54	◒	●	○	●	●	—	—	3	—	d	—
Clock radios with cassette player														
General Electric 7-4965	60/50	10x30x15	85	◑	◑	◑	◑	◑	✔	✔	12	A,D,G,H.K,O,S,T	—	B,E
Panasonic RC-X250	[1]	10x33x13	76	◑	◑	○	◑	◑	✔	✔	12	A,G,K,O,R,U	b,c	A,H
Sony ICF-CS650	75/65	15x28x15	74	○	◑	◓	○	◒	✔	✔	12	G,R,T,U	c,f,i	A,F,H
Soundesign 3844MGY	40/30	13x30x13	62	○	●	●	●	◒	—	—	3	G,K,J,S,U		F,G,I,M

[1] Discontinued. Replaced by RC-X260, $79 list and $60 average advertised sale price.

Features in Common
All: • Permit snooze time of about 8 min. • Retain time settings during short power failures.
Except as noted, all have: • Battery backup for clock and alarm memory. • Red display digits 1 cm. high. • Sleep-time radio play for up to 60 min. before automatic shutoff. • Switch to reset alarm.

Keys to Advantages
A–Alarm works despite power failure.
B–Shows actual time plus up to 2 alarm times.
C–Twin alarms settable for 2 different stations.
D–Tone alarm has adjustable volume control.
E–Memory needs no battery.
F–Digital tuner with presettable stations.
G–Tuner can receive in stereo.
H–Battery-strength indicator.
I–Illuminated tuning dial.
J–Illuminated tuning pointer.

K–Earphone jack.
L–Nap timer.
M–Audio input for tape deck or CD player.
N–Display can show date and time.
O–Display has high/low brightness switch.
P–Display has larger digits than most.
Q–Night light—adjusts for room light.
R–Bass-boost tone control.
S–Treble-cut tone control.
T–Better than most in tuning ease.
U–Better than most in image rejection.

Key to Disadvantages
a–Possible to reset time by accident.
b–Controls for time-setting or dimmer inconveniently located on radio's bottom or rear.
c–Display dimmer than most in brightly lit room.
d–Radio volume must be turned completely down for alarm buzzer to sound.
e–Lacks alarm buzzer; radio is sole alarm.
f–Lacks indication alarm is set.
g–Lacks alarm-reset button.
h–Time-setting lacks fast reverse.
i–No slow forward, fast reverse for time setting.

Key to Comments
A–Display shows green digits.
B–Display shows blue digits.
C–Display uses LCD (liquid crystal) digits.
D–Terminals for external antenna.
E–3-position graphic equalizer.
F–Cassette player lacks Record function.
G–Cassette player lacks Rewind function.
H–Model permits wake-up to cassette play.
I–Cassette-deck flutter worse than most.
J–Warranty repairs cost $3 for handling.
K–Warranty repairs cost $3.50 for handling.
L–Warranty repairs cost $6 for handling.
M–Warranty repairs cost $10 for handling.

INTERPRETING THE QUANTITATIVE LITERACY LEVELS

Since adults are frequently required to perform arithmetic operations in everyday life, the ability to perform quantitative literacy tasks is another important aspect of literacy. These skills may seem, at first glance, to be fundamentally different from the types of knowledge and skill associated with prose and document literacy and therefore, to extend the concept of literacy beyond its traditional limits. However, experience in North America with large-scale assessments of adults indicates that the processing of printed information plays an important role in affecting the difficulty of tasks along the scale (Kirsch et al. 1993; Montigny et al. 1991).

In general, it appears that many individuals can perform single arithmetic operations when both the numbers and operations are made explicit. However, when the numbers to be used must be located in and extracted from different types of documents that contain similar but irrelevant information, when the operations to be used must be inferred from printed directions, and when multiple operations must be performed, the tasks become increasingly difficult.

The IALS quantitative literacy scale contains 33 tasks ranging from 229 to 408 in difficulty. These tasks are distributed as follows: Level 1 (1 task); Level 2 (9 tasks); Level 3 (16 tasks); Level 4 (5 tasks); and Level 5 (2 tasks). The difficulty of these tasks and, therefore, their placement along the scale, appears to be a function of several factors including:

- the particular arithmetic operation required to complete the task;
- the number of operations needed to perform the task successfully;
- the extent to which the numbers are embedded in printed materials; and
- the extent to which an inference must be made to identify the type of operation to be performed.

A detailed discussion of the five levels of quantitative literacy follows.

Quantitative Level 1 — Score range: 0 to 225

Although no quantitative tasks used in the IALS fall below the score value of 225, experience suggests that such tasks would require the reader to perform a single, relatively simple operation (usually addition) for which either the numbers are already entered onto the given document and the operation is stipulated, or the numbers are provided and the operation does not require the reader to borrow.

The easiest quantitative task in the IALS (225) directs the reader to complete an order form. The last line on this form says 'Total with Handling.' The line above it says 'Handling Charge $2.00.' The reader simply had to add the $2.00 to the $50.00 they had entered on a previous line to indicate the cost of the tickets. In this task, one of the numbers was stipulated, the operation was easily identified from the word 'total' and the operation did not require the reader to borrow or carry numbers. Moreover, the format of the form set the problem up in a simple column format, further facilitating the task for the reader.

Quantitative Level 2 — Score range: 226 to 275

Tasks in this level typically require readers to perform a single arithmetic operation (frequently addition or subtraction) using numbers that are easily located in the text or document. The operation to be performed may be easily inferred from the wording of the question or the format of the material (for example, a bank deposit form or an order form).

A typical Level 2 task on the quantitative scale directs the reader to use a weather chart in a newspaper to determine how many degrees warmer today's high temperature is expected to be in Bangkok than in Seoul. Here the reader had to cycle through the table to locate the two temperatures and then subtract them to determine the difference. This task received a difficulty value of 255.
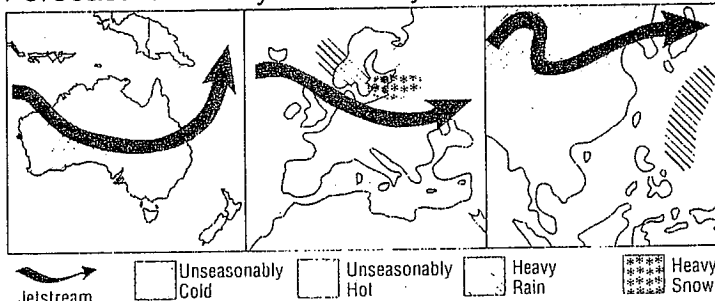
WEATHER

Forecast for Friday to Sunday

Europe

	Today High C	Low C	W	Tomorrow High C	Low C	W
Algarve	19	7	s	21	9	s
Amsterdam	11	6	pc	12	7	pc
Ankara	17	7	pc	19	8	pc
Athens	22	15	pc	23	14	pc
Barcelona	16	8	s	14	9	s
Belgrade	14	6	pc	10	1	c
Berlin	8	2	c	6	1	c
Brussels	11	6	pc	14	7	pc
Budapest	9	1	pc	9	2	c
Copenhagen	7	1	i	6	2	c
Costa del Sol	21	8	s	13	10	s
Dublin	10	6	pc	13	8	pc
Edinburgh	10	6	c	10	6	c
Florence	11	5	s	14	6	s
Frankfurt	12	6	pc	13	4	pc
Geneva	9	2	s	12	4	s
Helsinki	-1	-7	sf	-3	-10	pc
Istanbul	17	10	pc	15	9	sh
Las Palmas	26	18	pc	27	18	pc
Lisbon	19	9	s	19	10	.s
London	12	5	pc	13	7	pc
Madrid	17	3	s	18	4	s
Milan	9	3	s	13	.6	s
Moscow	1	-3	r	-3	-11	sf
Munich	11	3	pc	12	6	pc
Nice	14	7	s	15	8	s
Oslo	4	-4	c	5	-2	c
Paris	12	6	pc	13	6	pc
Prague	11	1	pc	8	2	c
Reykjavik	4	2	r	6	-1	c
Rome	20	12	s	20	10	s
St. Petersburg	-1	-7	sf	-4	-12	pc
Stockholm	1	-5	sn	-2	-7	c
Strasbourg	12	5	pc	15	7	pc
Tallinn	-1	-7	sf	-4	-10	pc
Venice	10	3	s	11	4	s
Vienna	9	-1	pc	10	2	c
Warsaw	8	2	sh	6	1	c
Zurich	8	0	s	9	1	pc

Middle East

	Today High C	Low C	W	Tomorrow High C	Low C	W
Beirut	28	19	pc	29	20	s
Cairo	29	20	pc	28	19	pc

Asia

	Today High C	Low C	W	Tomorrow High C	Low C
Bangkok	32	22	pc	30	23
Beijing	11	0	s	8	2
Hong Kong	30	23	s	29	22
Manila	31	25	s	31	25
New Delhi	31	13	s	32	16
Seoul	14	6	pc	14	4
Shanghai	22	10	pc	24	12
Singapore	31	24	pc	28	23
Taipei	26	21	pc	26	19
Tokyo	18	9	pc	17	7

Africa

Algiers	27	14	s	26	13
Cape Town	20	11	sh	18	11
Casablanca	20	14	c	21	11
Harare	34	17	s	32	18
Lagos	30	24	t	29	24
Nairobi	27	12	pc	26	13
Tunis	27	17	pc	17	14

North America

Anchorage	0	-2	c	3	0
Atlanta	14	4	pc	8	2
Boston	15	4	c	8	-1
Chicago	2	-5	c	-2	-8
Denver	8	-3	pc	4	-6
Detroit	4	-2	c	4	-5
Honolulu	31	20	s	31	21
Houston	15	3	pc	12	6
Los Angeles	28	14	s	24	13
Miami	30	22	pc	29	21
Minneapolis	-1	-8	c	1	-7
Montreal	7	-2	sl	4	-3
Nassau	31	22	pc	28	21
New York	14	4	r	10	2
Phoenix	23	11	pc	22	8
San Fran.	20	11	pc	21	8
Seattle	11	6	pc	13	7
Toronto	6	-3	c	3	-3
Washington	14	6	r	11	4

Jetstream — Unseasonably Cold — Unseasonably Hot — Heavy Rain — Heavy Snow

Australia
Warm weather will engulf the southeast coast of Australia Friday and over the weekend. Although it will be warm in Melbourne, Sydney and Adelaide, showers are expected. Hobart will have some sunshine and warmer temperatures each day.

Europe
Western and central Europe will have a spell of mild weather Friday into the weekend. London and Paris will have dry weather with some sunshine Friday into Sunday. Rain will continue to soak southwestern Norway. Snow will blanket the area from Minsk to Moscow.

Asia
Typhoon Elsie will probably stay to the east of the Philippines and south of Japan Friday and the weekend. Some rain is apt to fall in Seoul and there could even be a little ice or snow. Cold air will pour into Beijing and snow is a possibility. Hong Kong will start the weekend warm.

Oceania

	Today High C	Low C	W	Tomorrow High C	Low C	W
Auckland	20	14	s	17	11	sh
Brisbane	29	22	s	31	23	pc
Canberra	25	12	s	26	14	s
Melbourne	23	17	sh	21	15	s
Sydney	27	17	pc	25	16	pc

Latin America

	Today High C	Low C	W	Tomorrow High C	Low C	W
Buenos Aires	23	11	pc	26	13	s
Caracas	29	20	s	31	18	s
Lima	23	17	c	23	16	c
Mexico City	23	11	sh	23	12	pc
Rio de Janeiro	32	22	s	28	21	sh
Santiago	24	4	s	22	6	pc

Legend: s-sunny, pc-partly cloudy, c-cloudy, sh-showers, t-thunderstorms, r-rain, sf-snow flurries, sn-snow, i-ice, W-Weather.

A similar but slightly more difficult task (268) requires the reader to use the chart about women in the teaching profession in Europe that is displayed in Level 1 for the document scale. This task directs the reader to calculate the percentage of men in the teaching profession in Italy. Both this task and the one just mentioned involved calculating the difference between two numbers. Part of what distinguishes these two tasks is that in the former, both temperatures could be identified in the table from the newspaper. For the task involving men teachers in Italy, the reader needed to make the inference that the percentage of men teachers is equal to 100% minus the percentage of women teachers.

Quantitative Level 3 — Score range: 276 to 325

Tasks found in this level typically require the reader to perform a single operation. However, the operations become more varied — some multiplication and division tasks are found in this level. Sometimes two or more numbers are needed to solve the problem and the numbers are frequently embedded in more complex displays. While semantic relation terms such as 'how many' or 'calculate the difference' are often used, some of the tasks require the reader to make higher order inferences to determine the appropriate operation.

For example, one task located at 302 on the quantitative scale directs the reader to look at two graphs containing information about consumers and producers of primary energy. In one question, they are asked to calculate how much more energy Canada produces than it consumes. Here the operation is not facilitated by the format of the document and the reader must locate the information using both bar graphs. In another question using this document, the reader is directed to calculate the total amount of energy in 'quadrillion Btu' consumed by Canada, Mexico and the United States of America. This task falls at 300 on the scale. It requires the reader to add three numbers. Presenting two graphs likely contributed to the difficulty of this task. Some respondents may have performed the appropriate calculation for the three countries specified using the producer energy chart rather than the consumer energy chart.

Another task at this level involves the fireworks chart referred to earlier for the document scale. This quantitative task asks the reader to calculate how many more people were injured in 1989 than in 1988. What contributes to this task receiving a difficulty value of 293 is that one of the numbers was not given in the line graph. The reader needed to interpolate the number from information provided along the vertical axis.

In a task falling in the same Level (located at 280 on the scale), readers are asked to look at a recipe for scrambled eggs with tomatoes. The recipe gives the ingredients for four servings: 3 tablespoons of oil, 1 garlic clove, 1 teaspoon of sugar, 500 grams of fresh red tomatoes and 6 eggs. The question asks them to determine the number of eggs they will need if they are using the recipe for six people. Here they must know how to calculate or determine the ratio needed. This task is somewhat easier than might be expected, given other tasks at this level. This may be because people are familiar with recipes and with manipulating them to fit a particular situation.

World's Major Producers of Primary Energy, 1990

	0	10	20	30	40	50	60	70	80	90
U.S.S.R.								67.64		
United States								67.47		
China				30.23						
Saudi Arabia		15.88								
Canada		13.15								
United Kingdom	8.66									
Iran	7.65									
Mexico	7.36									
India	6.75									
Norway	6.13									
West Germany	5.77									
Indonesia	4.92									

SOURCE: Energy Information Administration; International Energy Annual 1990; Quadrillion (10^{15}) Btu

World's Major Consumers of Primary Energy, 1990

	0	10	20	30	40	50	60	70	80	90
United States									81.17	
U.S.S.R.							57.15			
China			28.85							
Japan		18.18								
West Germany	12.47									
Canada	10.79									
United Kingdom	9.13									
France	8.69									
India	7.99									
Brazil	5.65									
Mexico	4.94									
Netherlands	3.39									

SOURCE: Energy Information Administration; International Energy Annual 1990; Quadrillion (10^{15}) Btu

This appears to be true for another question using this recipe. It asks the reader to determine the amount of oil that would be needed if the recipe were being used for two people. This task received a value of 253 on the scale. A larger percentage of respondents found it easier to halve an ingredient than to increase one by 50%. It is not clear why this is so. It may be that some of the respondents have an algorithm for responding to certain familiar tasks that does not require them to apply general arithmetic principles for solving the problem.

Quantitative Level 4 — Score range: 326 to 375

With one exception, the tasks at this level require the reader to perform a single arithmetic operation where typically either the quantities or the operation are not easily determined. That is, for most of the tasks at this level, the question or directive does not provide a semantic relation term such as 'how many' or 'calculate the difference' to help the reader.

One such task involves a compound interest table. It directs the reader to 'calculate the total amount of money you will have if you invest $100 at a rate of 6% for 10 years.' This task received a difficulty value of 348, in part because many people treated this as a document rather than a quantitative task and simply looked up the amount of interest that would be earned. They may have forgotten to add the interest to their $100 investment.

Compound Interest
Compounded Annually

Principal	Period	4%	5%	6%	7%	8%	9%	10%	12%	14%	16%
$100	1 day	0.011	0.014	0.016	0.019	0.022	0.025	0.027	0.033	0.038	0.044
	1 week	0.077	0.096	0.115	0.134	0.153	0.173	0.192	0.230	0.268	0.307
	6 months	2.00	2.50	3.00	3.50	4.00	4.50	5.00	6.00	7.00	8.00
	1 year	4.00	5.00	6.00	7.00	8.00	9.00	10.00	12.00	14.00	16.00
	2 years	8.16	10.25	12.36	14.49	16.64	18.81	21.00	25.44	29.96	34.56
	3 years	12.49	15.76	19.10	22.50	25.97	29.50	33.10	40.49	48.15	56.09
	4 years	16.99	21.55	26.25	31.08	36.05	41.16	46.41	57.35	68.90	81.06
	5 years	21.67	27.63	33.82	40.26	46.93	53.86	61.05	76.23	92.54	110.03
	6 years	26.53	34.01	41.85	50.07	58.69	67.71	77.16	97.38	119.50	143.64
	7 years	31.59	40.71	50.36	60.58	71.38	82.80	94.87	121.07	150.23	182.62
	8 years	36.86	47.75	59.38	71.82	85.09	99.26	114.36	147.60	185.26	227.84
	9 years	42.33	55.13	68.95	83.85	99.90	117.19	135.79	177.31	225.19	280.30
	10 years	48.02	62.89	79.08	96.72	115.89	136.74	159.37	210.58	270.72	341.14
	12 years	60.10	79.59	101.22	125.22	151.82	181.27	213.84	289.60	381.79	493.60
	15 years	80.09	107.89	139.66	175.90	217.22	264.25	317.72	447.36	613.79	826.55
	20 years	119.11	165.33	220.71	286.97	366.10	460.44	572.75	864.63	1,274.35	1,846.08

Another task at this level requires respondents to read a newspaper article describing a research finding linking allergies to a particular genetic mutation. The question directs the reader to calculate the number of people studied who were found to have the mutant gene. To answer the question correctly, readers must know how to convert the phrase '64 percent' to a decimal number and then multiply it by the number of patients studied (400). The text provides no clues on how to set up this problem.

A third task involves the distance chart shown below. Readers were asked to 'calculate the total number of kilometres travelled in a trip from Guadalajara to Tecoman and then to Zamora.' Here a semantic relation term was provided, but the quantities were not easily identified. As a result, this task received a difficulty value of 335. Making the inference that the trip was from Guadalajara to Tecoman and then from Tecoman to Zamora was difficult for some respondents. In a different task, respondents were asked to determine how much less the distance from Guadalajara to Tecoman is than the distance from Guadalajara to Puerto Vallarta. In this Level 3 task (308), the quantities were relatively easy to locate.

TABLE OF APPROXIMATE DISTANCES (in kilometres)

Quantitative Level 5 — Score range: 376 to 500

These tasks require readers to perform multiple operations sequentially, and they must disembed the features of the problem from the material provided or rely on background knowledge to determine the quantities or operations needed.

One of the most difficult tasks on the quantitative scale (381) requires readers to look at a table providing nutritional analysis of food and then, using the information given, determine the percentage of calories in a Big Mac® that comes from total fat. To answer this question, readers must first recognise that the information about total fat provided is given in grams. In the question, they are told that a gram of fat has 9 calories. Therefore, they must convert the number of fat grams to calories. Then, they need to calculate this number of calories as a percentage of the total calories given for a Big Mac®. Only one other item on this scale received a higher score.

Nutritional Analysis

	Serving Size	Calories	Protein (g)	Carbohydrates (g)	Total Fat (g)	Saturated Fat (g)	Monounsaturated Fat (g)	Polyunsaturated Fat (g)	Cholesterol (mg)	Sodium (mg)
Sandwiches										
Hamburger	102 g	255	12	30	9	5	1	3	37	490
Cheeseburger	116 g	305	15	30	13	7	1	5	50	725
Quarter Pounder®	166 g	410	23	34	20	11	1	8	85	645
Quarter Pounder® w/Cheese	194 g	510	28	34	28	16	1	11	115	1110
McLean Deluxe™	206 g	320	22	35	10	5	1	4	60	670
McLean Deluxe™ w/Cheese	219 g	370	24	35	14	8	1	5	75	890
Big Mac®	215 g	500	25	42	26	16	1	9	100	890
Filet-O-Fish®	141 g	370	14	38	18	8	6	4	50	730
McChicken®	187 g	415	19	39	19	9	7	4	50	830
French Fries										
Small French Fries	68 g	220	3	26	12	8	1	2.5	0	110
Medium French Fries	97 g	320	4	36	17	12	1.5	3.5	0	150
Large French Fries	122 g	400	6	46	22	15	2	5	0	200
Salads										
Chef Salad	265 g	170	17	8	9	4	1	4	111	400
Garden Salad	189 g	50	4	6	2	1	0.4	0.6	65	70
Chunky Chicken Salad	255 g	150	25	7	4	2	1	1	78	230
Side Salad	106 g	30	2	4	1	0.5	0.2	0.3	33	35
Croutons	11 g	50	1	7	2	1.3	0.1	0.5	0	140
Bacon Bits	3 g	15	1	0	1	0.3	0.2	0.5	1	95

Soft Drinks

	Coca-Cola Classic®				diet Coke®				Sprite®			
	Small	Medium	Large	Jumbo	Small	Medium	Large	Jumbo	Small	Medium	Large	Jumbo
Calories	140	190	260	380	1	1	2	3	140	190	260	380
Carbohydrates (g)	38	50	70	101	0.3	0.4	0.5	0.6	36	48	66	96
Sodium (mg)	15	20	25	40	30	40	60	80	15	20	25	40

ESTIMATING LITERACY PERFORMANCE ACROSS THE LEVELS

The literacy levels not only provide a means for exploring the progression of information-processing demands across each of the scales, but they also can be used to help explain how the proficiencies individuals demonstrate reflect the likelihood they will respond correctly to the broad range of tasks used in this assessment as well as to similar tasks that were not included. In practical terms, this means that individuals performing at 250 on each scale are expected to be able to perform the average Level 1 and 2 tasks with a high degree of proficiency. That is, they will be able to perform these kinds of tasks with an average probability of 80% or higher. It does not mean that they will not be able to perform tasks in Levels 3 or higher. They will do so some of the time, but not consistently.

Table 1 displays the probability that individuals performing at selected points on each of the scales will give a correct response to tasks of varying difficulty. For example, a reader whose prose proficiency is 150 has less than a 50% chance of giving a correct response to the Level 1 tasks. Individuals whose proficiency score is 200, in contrast, have about an 80% probability of responding correctly to these Level 1 tasks. In terms of task demands, it can be inferred that adults performing at 200 on the prose scale are likely to be able to locate a single piece of information in a brief text when there is no distracting information, or if plausible but incorrect information is present but located away from the correct answer. However, these individuals are likely to demonstrate far more difficulty with tasks in Levels 2 through 5. For example, they would have only a 40% chance of performing the average Level 2 task correctly and an 18% chance of success with tasks in Level 3 and no more than a 7% chance with tasks in Levels 4 and 5.

In contrast, respondents demonstrating a proficiency of 300 on the prose scale have about an 80% chance or higher of succeeding on tasks in Levels 1, 2 and 3. This means that they demonstrate success with tasks that require them to make low-level inferences and with tasks that require them to take some conditional information into account. They can also integrate or compare and contrast information that is easily identified in the text. On the other hand, they are likely to demonstrate some difficulty with tasks where they must make high text-based inferences or where they need to process more abstract types of information. These more difficult tasks may also require them to draw on less familiar or more specialised types of knowledge beyond that given in the text. On average, they have about a 50% probability of performing Level 4 tasks correctly; with Level 5 tasks, their likelihood of responding correctly decreases to 40%.

ES

TABLE 1 SELECTED PROFICIENCY SCORES, By Literacy Scale Levels

SELECTED PROFICIENCY SCORES......................

	150	200	250	300	350
	%	%	%	%	%
PROSE SCALE					
Level 1	48	81	95	99	100
Level 2	14	40	76	94	99
Level 3	6	18	46	78	93
Level 4	2	7	21	50	80
Level 5 (a)	2	6	18	40	68
DOCUMENT SCALE					
Level 1	40	72	94	99	100
Level 2	20	51	82	95	99
Level 3	7	21	50	80	94
Level 4	4	13	34	64	85
Level 5(a)	<1	1	3	13	41
QUANTITATIVE SCALE					
Level 1(a)	34	67	89	97	99
Level 2	21	47	76	92	98
Level 3	7	21	51	81	94
Level 4	1	6	22	57	86
Level 5(a)	1	2	7	20	53

(a) Based on one task.

Similar kinds of interpretations can be made using the information presented for the document and quantitative literacy scales. For example, someone who is at 200 on the quantitative scale has, on average, a 67% chance of responding correctly to Level 1 tasks. His or her likelihood of responding correctly decreases to 47% for Level 2 tasks, 21% for Level 3 tasks, 6% for Level 4 tasks and only 2% for Level 5 tasks. Similarly, readers with a proficiency of 300 on the quantitative scale would have a probability of 92% or higher of responding correctly to tasks in Levels 1 and 2. Their average probability would decrease to 81% for Level 3 tasks, 57% for Level 4 and 20% for Level 5.

ESTIMATING THE VARIABILITY OF LITERACY TASKS ACROSS COUNTRIES

One of the goals in conducting international surveys is to be able to compare populations on common scales. In this study, three literacy scales were used to compare both the distributions of literacy skills and the relationships between literacy skills and a variety of social, educational and labor-market variables. Each literacy scale consisted of more than 30 literacy tasks which received item parameters that define the difficulty of the task and how well it discriminates among the populations of adults who responded to the task. These parameters were determined based on how adults within and across participating countries responded to each task.

Table 1 shows the average probabilities of successful performance by individuals with selected proficiency scores on tasks in each literacy level of the prose, document and quantitative scales.

Under standard assumptions of item response theory (IRT), item parameters are thought to be invariant across respondents and across subpopulations. However, we have discovered through the conduct of large-scale assessments that this assumption is not always true. In IALS, we have noted in the technical report (Yamamoto 1997) that some language/country subpopulations do respond differently to a subset of literacy tasks. As described in the technical report, individual items were dropped from the assessment if at least seven of the original 10 language/country populations were shown not to have the same item parameters. That is, if the response data for a particular item showed poor fit to the item parameters common to the rest of the language/country populations. In addition, if there were items in which only 1, 2 or 3 countries varied, these countries were allowed to have unique parameters for that item. This resulted in a total of 13 items being dropped from the assessment with 31 items getting a unique parameter for one language/country population, 16 for two language/country populations, and six for three language/country populations. Another way to look at this is that there were a total of 1,010 constraints (114 items minus the 13 dropped times 10 language samples). Of these, unique item parameters were required or allowed in 81 instances meaning that 92% of the constraints support a common scale across the 10 original language/country populations. France was eventually removed from the analyses at their request.

The reasons for these discrepancies were due largely to differences in translations among countries or to differences in interpretation of scoring rubrics for individual items. The differential performance on some items also reflected the variation in language and culture although no obvious or specific reason could be identified. The fact that not all items had identical item parameters resulted in two types of variation. The first is that the differences could influence the distribution of proficiency scores for a particular country/language group. Analyses indicated that the consequence of using a partially different set of item parameters on the proficiency distribution for a particular population was minimal. For any given population, when the proficiency distribution was estimated based on either a set of items which included those common across countries as well as those unique to a given country or on a set of items which were optimal for a different population, the means and standard deviations of estimated proficiencies differed by less than half of a standard error. Typically, standard errors of estimation ranged between one and three points on the 500-point scales depending on a particular language/country population.

Another type of variation which results from having a small set of items with unique parameters is the placement of particular tasks along the scales according to their 'RP80' values. The 'RP80' value is the value on the scale of 0–500 which indicates the difficulty of a particular task — the higher the value, the more difficult the task. At the beginning of this appendix, we discussed the fact that a criterion of 80 percent was used meaning that tasks were placed along a scale based on the probability that someone with that level of proficiency would have an 80 percent chance of getting that task and others like it correct. The fact that a small subset of tasks have unique parameters for particular county/language groups means that some tasks fall at different points along each scale. Since this appendix is about describing what it means to be at a particular point along each literacy scale and uses exemplar tasks to reflect on this meaning, it seems important to try to describe the extent of the variation which exists among the countries with respect to the placement of tasks along the literacy scales.

To evaluate the variability of RP80s for each language/country population, the deviation of RP80s against the common RP80 was examined. It is important to note that no country received all common item parameters. That is, at least one item for each country received a unique set of item parameters. However, at least seven of the original language/country populations received common parameters for each of the 101 items. In total, there are 15 country/language groups for which we have data to estimate this variation. Nine of the groups are from the first assessment and six are from the assessment just completed. There were a total of 101 literacy exercises meaning that there could be as many as 1,515 deviations (101 times 15). The mean deviation among the RP80s was 4.7 with a standard deviation of 14.1. This means that the average variation among the RP80s for the literacy tasks was 4.7 points on a 500-point scale or less than 10% of the 50 points making up a particular literacy level. In addition, a small number of items had large deviations which made up a large percentage of this variation. Only 1% of the actual deviations observed accounts for about 20% of the average deviation. That is, 99% of the deviations have a mean of 3.6, or a 20% reduction from the average of 4.7. Table 2 shows the average deviation of RP80s by each of the 15 language/country groups. Here we can see that the average deviation ranges from a low of 1.1 for the Swiss/French to 7.6 for Australia.

TABLE 2 AVERAGE DEVIATION OF RP80 VALUES, By Country/Language

Country (language group)	Average deviation of RP80 values
Australia	7.6
Belgium	5.8
Canada (English)	3.6
Canada (French)	3.2
Germany	5.3
Ireland	4.5
Netherlands	3.4
New Zealand	7.2
Northern Ireland	6.9
Poland	5.4
Sweden	5.2
Switzerland (French)	1.1
Switzerland (German)	4.0
Great Britain	5.2
United States of America	2.0

CONCLUSION

One of the goals of large-scale surveys is to provide a set of information that can inform policy makers and help them during the decision-making process. Presenting information in a way that will enhance understanding of what has been measured and what conclusions may be drawn from the data is important to reaching this goal. This appendix has presented a framework for understanding the consistency of task responses demonstrated by adults from a number of countries. This framework identifies a set of variables shown to underlie successful performance on a broad array of literacy tasks. Collectively, these variables provide a means for moving away from interpreting survey results in terms of discrete tasks or a single number and towards identifying levels of performance that have generalisability and validity across assessments and groups.

The knowledge and understanding such a framework provides contribute to the evolving concept of test design as more than merely assigning a numerical value (or position) to an individual based on his or her responses to a set of tasks, but rather to assigning meaning and interpretability to this number. As concern ceases to centre on discrete behaviours or isolated observations and concentrates more on providing a meaningful score, a higher level of measurement is reached (Messick 1989).