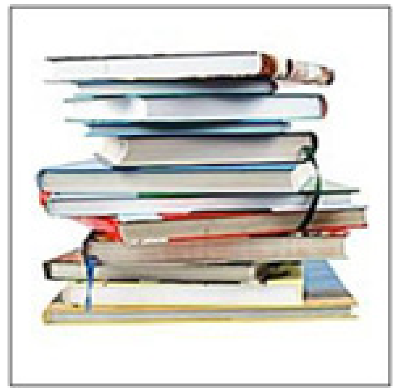
User Guide

Public Use Microdata File (PUMF)



The Programme for the International Assessment of Adult Competencies, 2012

Tourism and the Centre for Education Statistics Division
Main Building, Room 2001, Ottawa, K1A 0T6
Telephone: 1-800-263-1136

Statistics Canada Statistique Canada

Canada

How to obtain more information

For information about this product or the wide range of services and data available from Statistics Canada, visit our website, www.statcan.gc.ca.

You can also contact us by

email at infostats@statcan.gc.ca,

telephone, from Monday to Friday, 8:30 a.m. to 4:30 p.m., at the following toll-free numbers:

• Statistical Information Service	1-800-263-1136
• National telecommunications device for the hearing impaired	1-800-363-7629
• Fax line	1-877-287-4369

Depository Services Program

• Inquiries line	1-800-635-7943
• Fax line	1-800-565-7757

Standards of service to the public

Statistics Canada is committed to serving its clients in a prompt, reliable and courteous manner. To this end, Statistics Canada has developed standards of service that its employees observe. To obtain a copy of these service standards, please contact Statistics Canada toll-free at 1-800-263-1136. The service standards are also published on www.statcan.gc.ca under "About us" > "The agency" > "Providing services to Canadians."

Published by authority of the Minister responsible for Statistics Canada

Cette publication est aussi disponible en français.

Note of appreciation

Canada owes the success of its statistical system to a long-standing partnership between Statistics Canada, the citizens of Canada, its businesses, governments and other institutions. Accurate and timely statistical information could not be produced without their continued co-operation and goodwill.

Standard symbols

The following symbols are used in Statistics Canada publications:

.	not available for any reference period
..	not available for a specific reference period
...	not applicable
0	true zero or a value rounded to zero
0^s	value rounded to 0 (zero) where there is a meaningful distinction between true zero and the value that was rounded
^{p}	preliminary
^{r}	revised
x	suppressed to meet the confidentiality requirements of the *Statistics Act*
^{E}	use with caution
F	too unreliable to be published
*	significantly different from reference category ($p < 0.05$)

Table of Contents

1 Introduction

What is the Programme for the International Assessment of Adult Competencies (PIAAC)?

The Programme for the International Assessment of Adult Competencies (PIAAC) study is part of an international program. It has been designed to see what activities adults perform in their daily lives, such as reading, finding information, and using computers and technology, and to learn about their education, work experience and use of key work skills on the job. The Organisation for Economic Co-operation and Development (OECD) PIAAC assessed the proficiency of adults in three information-processing skills essential for full participation in the knowledge-based economies and societies of the 21st century: literacy, numeracy and problem solving in technology-rich environments.

> *PIAAC is a multi-cycle international programme of assessment of adult skills and competencies sponsored by the Organisation for Economic Co-operation and Development (OECD). Governments in OECD countries face the challenge of maintaining competitiveness in a global economy by ensuring that labour markets are flexible and responsive and are open to a wide range of people of all ages.*
>
> *PIAAC assesses the level and distribution of adult skills across countries, focusing on the cognitive and workplace skills needed for successful participation in the economy and society of the 21st century. PIAAC collects information on skills required in the workplace, participants' educational backgrounds and professional attainments, and their ability to sue information and communications technology. In addition, PIAAC includes an assessment of cognitive skills to measure participants' general levels of literacy and numeracy.*

The design and implementation of PIAAC is the responsibility of an international consortium of well established institutions from North America and Europe led by the Educational Testing Service in the United States. The other partners of this consortium are Westat in the United States; cApStAn in Belgium; the Research Centre for Education and the Labour Market (ROA) at the University of Maastricht in the Netherlands; and the GESIS - Leibniz Institute for the Social Sciences, the German Institute for International Education Research (DIPF), and the Data Processing Centre of the International Association for the Evaluation of Educational Achievement (IEA) in Germany.

The PIAAC builds on Canada's past experiences from similar studies conducted in 1994 (International Adult Literacy Survey – IALS) and 2003 (International Adult Literacy and Skills Survey – IALSS).

The PIAAC required all participating countries[1] to collect data from a nationally representative sample of at least 5,000[2] respondents aged 16 to 65 for each language tested – English and French in the case of Canada. The minimum sample requirements for the PIAAC survey were exceeded in the Canadian PIAAC due to contributions by a number of federal departments and provincial governments that allowed for the collection of additional records, so as to ensure high reliability in the estimation of data values for sub-

[1] Participating Countries: Australia, Austria, Belgium – Flanders, Canada, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Ireland, Italy, Japan, Korea, Netherlands (The), Norway, Poland, Russian Federation, Slovak Republic, Spain, Sweden, United Kingdom (England), United Kingdom (Northern Ireland), United States.

[2] Minimum was 4,500 if the problem-solving component was not conducted.

national populations. As a result, the number of respondents on the full dataset is sufficient to provide accurate estimates for the Yukon, Northwest Territories and Nunavut. Finally, the 2012 PIAAC added the following supplementary sub-populations to the sample: recent immigrants, aboriginals/Métis, youth (persons aged 16-25) and linguistic minorities (Anglophones in Québec, Francophone in Ontario, New Brunswick, and Manitoba).

Data collection in Canada began in November 2011 and ended in June 2012. By the end of these eight months, over 27,000 individuals from across Canada had spent an average of one hour and forty-five minutes responding to the PIAAC and contributing to the value of the survey. Table 1.1 shows the actual and weighted distributions of respondents from across Canada.

Table 1.1 Geographical distribution of PIAAC respondents, Canada, provinces and territories, 16-65 years, 2012

Province or Territory	**Number of respondents**	**Representative population (16-65)**
Newfoundland and Labrador	1,609	349,233
Prince Edward Island	929	97,542
Nova Scotia	1,441	627,538
New Brunswick	1,686	500,997
Quebec	5,911	5,404,254
Ontario	5,313	9,148,632
Manitoba	2,312	785,291
Saskatchewan	1,601	657,025
Alberta	1,224	2,622,199
British Columbia	2,733	3,111,300
Yukon Territory	830	25,564
Northwest Territory	917	30,506
Nunavut	779	20,987
Canada (Total)	**27,285**	**23,381,067**

References

OECD. 2011. *PIAAC Technical Standards and Guidelines.* December 2011. pp. 6-7. pp. 89-91. http://www.oecd.org/site/piaac/publications.htm

2 Background

PIAAC has evolved from two previous international literacy surveys: the International Adult Literacy Survey (IALS), conducted between 1994 and 1998, and the Adult Literacy and Lifeskills Survey (ALL), conducted between 2003 and 2006. The 1994 IALS was the first effort to undertake a large-scale assessment of adult literacy skills at the international level. IALS compared the demonstrated literacy skills of people across countries, cultures and languages. Thus, it highlighted the importance of literacy to the economic and social well-being of countries. Twenty-three countries participated in IALS: Australia, Belgium (Flanders), Canada, Chile, the Czech Republic, Denmark, Finland, France, Germany, Ireland, Italy, Mexico (Nuevo Leon), the Netherlands, New Zealand, Northern Ireland, Norway, Poland, Portugal, Slovenia, Sweden, Switzerland, the United Kingdom and the United States.

The practical experience gained from the implementation of IALS and from the analysis of the survey results indicated that there was room for improvement in future surveys of this type. In particular, the need was identified to standardize, as much as practically possible, the survey design, survey implementation procedures and estimation methods. ALL therefore sought to improve on IALS by developing standards to ensure that minimum quality assurance goals would be met, sources of survey variability could be minimized and survey results could be compared. Eleven countries participated in ALL: Australia, Bermuda, Canada, Hungary, Italy, Korea, Norway, the Netherlands, New Zealand, Switzerland and the United States.

PIAAC seeks to ensure continuity with these previous surveys, to extend the concept of literacy and numeracy to problem solving in technology-rich environments and to provide more information about individuals with low proficiency levels by assessing reading component skills. The standards developed for PIAAC are based on, and expand upon, the ALL standards.

References

OECD. 2011. *PIAAC Technical Standards and Guidelines.* December 2011. pp. 6-7. pp. 89-91.
http://www.oecd.org/site/piaac/publications.htm

3 Survey Overview

Every respondent to the PIAAC was first given a common questionnaire seeking information about demographic characteristics and variables such as educational attainment, occupation, income, and engagement in adult learning and community activities.

The respondents were then given an internationally validated psychometric[3] instrument designed to measure proficiency in four domains:

1. **Literacy** – Understanding, evaluating, using and engaging with written text.
2. **Numeracy** – The knowledge and skills required to effectively manage the mathematical demands of diverse situations.
3. **Problem Solving in Technology-Rich Environments –** Solving problems by setting up appropriate goals and plans, accessing and using information through computers.
4. **Reading Components –** Basic reading skills such as decoding, word recognition and sentence understanding.

The central element of the survey was the direct assessment of the literacy, numeracy and problem solving skills of respondents using commonplace tasks of varying degree of difficulty drawn from a range of topic and knowledge areas. This information was supported by the collection of background information on respondents. In addition, the background questionnaire included questions on the self-assessment of literacy and numeracy skills of respondents, on the training which the respondent has taken in the year previous to the survey and on the perceived barriers to realizing enhanced literacy or numeracy skill levels.

In-person interviews were used to complete the background questionnaire and to administer the direct assessments (i.e. literacy, numeracy, component skills and/or problem solving in technology-rich environments).

In-person interviews were used to select respondents. A computer-assisted data collection method was used at all stages of the data collection, including completion of the background questionnaire.

The direct assessments were available in paper- and computer-based formats. They were administered in the mode specified by the computer system that was used to collect the background questionnaire data.

Proxy responses were not acceptable for the background questionnaire or the direct assessments.

[3] Psychometrics refers to the branch of psychology that deals with the design, administration, and interpretation of quantitative tests for the measurement of variables such as intelligence, aptitude, and personality traits.

4 Survey Objectives

The primary objectives of PIAAC were to:

1) identify and measure cognitive competencies believed to underlie both personal and societal success,
2) assess the impact of these competencies on social and economic outcomes at individual and aggregate levels,
3) gauge the performance of education and training systems in generating required competencies,
4) help to clarify the policy levers that could contribute to enhancing competencies, and
5) link back to earlier assessments (i.e. IALS) of adult literacy and numeracy. Since much of the content is similar to that collected in previous surveys done in Canada and elsewhere in the world (International Adult Literacy Skills Survey – IALSS) in both 2003 and 1994, the PIAAC will allow researchers to measure how skills are changing over time.

One of PIAAC's core objectives was to assess how well participants use information and communications technology to access, manage, integrate and evaluate information; construct new knowledge; and communicate with other people. In addition, PIAAC collected information on participants' use of key work skills in their jobs, a first for an international study. In this way, PIAAC offers a far more complete and nuanced picture of the "human capital" on which countries can count as they compete in today's global economy. It helps policy makers assess the effectiveness of education and training systems, both for recent entrants to the labour market and for older people who may need to continue learning new skills throughout their lifetimes.

References

OECD. 2011. *PIAAC Technical Standards and Guidelines.* December 2011. pp. 6-7. pp. 89-91. http://www.oecd.org/site/piaac/publications.htm

5 Concepts and Definitions

This chapter is taken from the chapter 1 of the Reader's Companion, OECD 2013. Also note that, the psychometric items presented in this chapter are the international items. The Canadian items may differ slightly.

5.1 Proficiency as a continuum

The competencies assessed in the PIAAC are understood as involving a continuum of proficiency. Individuals are considered to be proficient to a greater or a lesser degree in the competency in question as opposed to being either "proficient" or "not proficient". In other words, there is no threshold that separates those who have the competency in question from those who do not. The measurement scales describe gradations in the complexity of the information-processing tasks in the domains of literacy, numeracy and problems solving in technology-rich environments. In each domain, this complexity is seen as a function of a small number of factors, such as the type of cognitive operations required by the task, the presence of distracting information, and the nature of information and knowledge required to successfully complete a task.

At the lower end of the proficiency scale, individuals have skills that allow them to undertake tasks of limited complexity, such as locating single pieces of information in short texts in the absence of other distracting information, or performing simple mathematical operations involving a single step, such as counting or ordering. At the highest level of proficiency, adults can undertake tasks that involve integrating information across multiple dense texts, reasoning by inference, working with mathematical arguments and models, and solving complex problems using information technologies that require navigation and the use of multiple tools.

Literacy and numeracy are often described as "basic" skills, in that they provide a "foundation" on which the development of other competencies rests. This description is unfortunate in that can give the impression that such skills are less complex than certain other "higher-order" skills or that the policy interest in such skills lies in ensuring that the population possesses an acceptable minimum or basic level of proficiency in these skills. It is important to emphasize that the objective of the PIAAC is to see how the adult population is distributed over the entire spectrum of proficiency in each of the domains assessed, not to assess whether adults have achieved a basic level of skills.

5.2 What the PIAAC measures

The PIAAC assesses the proficiency of adults in three information-processing skills essential for full participation in the knowledge-based economies and societies of the 21st century: literacy, numeracy and problem solving in technology-rich environments. This chapter describes the constructs measured in the survey and the information sought regarding skills use and the characteristics of respondents. First, a general description of the survey's approach to assessing adult skills is provided.

5.2.1 An assessment of key information-processing competencies

The skills assessed in the PIAAC are conceived as "key information-processing competencies". They represent skills essential for accessing, understanding, analyzing and using text-based information and, in the case of some mathematical information, information in the form of representations (e.g. pictures, graphs). These texts and representations may exist in the form of printed material or screen-based displays.

They are considered to be "key-information processing skills" in that they are:

- necessary for fully integrating and participating in the labour market, education and training, and social and civic life;
- highly transferable, in that they are relevant to many social contexts and work situations; and
- "learnable" and, therefore, subject to the influence of policy.

At the most fundamental level, literacy and numeracy constitute a foundation for developing higher-order cognitive skills, such as analytic reasoning, and are essential for gaining access to and understanding specific domains of knowledge. In addition, these skills are relevant across the range of life contexts, from education through work to home life and interaction with public authorities. In information-rich societies, in which information in text format (whether print-based or digital) is ubiquitous, a capacity to read and respond to text-based information is essential, whether that means understanding the user information on a packet of medicine, reacting appropriately to a memo from a colleague or superior at work, or enrolling a child at school. Similarly, numeracy skills are essential in most areas of life, from buying and selling goods, to understanding pension entitlements, to planning one's working day.

In addition, the capacity to manage information and solve problems in technology-rich environments – that is, to access, evaluate, analyze and communicate information – is becoming as important as understanding and interpreting text-based information and being able to handle mathematical content. Information and communication technology (ICT) applications have become a feature in most workplaces, in education, and in everyday life.

5.2.2 A use-oriented conception of competence

Literacy, numeracy and problem solving are competences that are essential for functioning in the modern world, for realizing the myriad tasks adults must undertake in the various life contexts. Adults read, deal with situations involving mathematical content and representations, and try to solve problems in order to do things and achieve certain objectives in a range of contexts. Consequently, the focus of the PIAAC is less on the mastery of certain content (e.g. vocabulary or arithmetical operations) and a set of cognitive strategies than on the ability to draw on this content and these strategies to successfully perform information-processing tasks in a variety of real-world situations.

5.3 Reporting the results

This section describes how the results from the PIAAC are reported. It discusses the proficiency scales and levels used to present the results of the assessment.

5.3.1 The proficiency scales

In each of the three domains of skills assessed, the results from the PIAAC are reported on a scale with a minimum value of 0 and a maximum of 500. Test takers and test items are located on the scale in terms of their proficiency and difficulty, respectively.

Individuals are located on the scale at the point at which they have a 67% chance of successfully completing items located at that point on the scale and at which an item has a probability of being correctly answered by 67% of individuals. An individual with a particular proficiency score will, on average, be able to complete items of lesser difficulty than those located at this score point with a greater chance of success, and be able to complete items of greater difficulty, but with less chance of success.

To illustrate this point, Table 5.1 shows the probability with which a person with a proficiency score of 300 on the literacy scale can successfully complete items of greater and lesser difficulty. As can be seen, a person with a proficiency score of 300 will successfully complete items of this level of difficulty 67% of the time, items with a difficulty value of 250, 95% of the time, and items with a difficulty value of 350, 28% of the time.

Table 5.1 Probability of successfully completing items of varying difficulty for a person scoring 300 on the literacy scale

	Proficiency score (literacy scale)			
	200	250	300	350
Probability of success	0.97	0.95	0.67	0.28

5.3.2 Proficiency levels

The proficiency scale in each of the domains assessed can be described in relation to the items that are located at the different points on the scale according to their difficulty.

With the exception of the lowest level (less than 1), tasks located at a particular level can be successfully completed approximately 50% of the time by a person with a proficiency score at the bottom of the range defining the level. In other words, a person with a score at the bottom of level 2 would score close to 50% in a test made up of items of level 2 difficulty. A person at the top of the level will get items located at that level correct most of the time. The "average" individual with a proficiency score in the range defining a level will successfully complete items located at that level approximately two thirds of the time.

5.3.3 Literacy and numeracy

Six proficiency levels are defined for the domains of literacy and numeracy; four levels are defined for problem solving in technology rich environments. The score-point ranges defining each level and the descriptors of the characteristics of tasks located at each of the levels can be found in Table 5.2. In the case of literacy and numeracy, the score-point ranges associated with each proficiency level are the same as those that apply in IALS and ALL for document and prose literacy and in ALL for numeracy. However, the descriptors that apply to the proficiency levels in the domains of literacy and numeracy differ between the PIAAC and IALS and ALL. This is because the domain of literacy in the OECD survey replaces the previously separate domains of prose and document literacy used in IALS and ALL, and because the OECD survey defines proficiency levels differently than the other surveys do.

Table 5.2 Proficiency levels: Literacy and numeracy

<table>
<tr><th>Level</th><th>Score range</th><th>Literacy</th><th>Numeracy</th></tr>
<tr><td>Below Level 1</td><td>0-175</td><td>The tasks at this level require the respondent to read brief texts on familiar topics to locate a single piece of specific information. There is seldom any competing information in the text and the requested information is identical in form to information in the question or directive. The respondent may be required to locate information in short continuous texts. However, in this case, the information can be located as if the text were non-continuous in format. Only basic vocabulary knowledge is required, and the reader is not required to understand the structure of sentences or paragraphs or make use of other text features. Tasks below Level 1 do not make use of any features specific to digital texts.</td><td>Tasks at this level require the respondents to carry out simple processes such as counting, sorting, performing basic arithmetic operations with whole numbers or money, or recognizing common spatial representations in concrete, familiar contexts where the mathematical content is explicit with little or no text or distractors.</td></tr>
<tr><td>1</td><td>176-225</td><td>Most of the tasks at this level require the respondent to read relatively short digital or print continuous, non-continuous, or mixed texts to locate a single piece of information that is identical to or synonymous with the information given in the question or directive. Some tasks, such as those involving non-continuous texts, may require the respondent to enter personal information onto a document. Little, if any, competing information is present. Some tasks may require simple cycling through more than one piece of information. Knowledge and skill in recognizing basic vocabulary determining the meaning of sentences, and reading paragraphs of text is expected.</td><td>Tasks at this level require the respondent to carry out basic mathematical processes in common, concrete contexts where the mathematical content is explicit with little text and minimal distractors. Tasks usually require one-step or simple processes involving counting; sorting; performing basic arithmetic operations; understanding simple percents such as 50%; and locating and identifying elements of simple or common graphical or spatial representations.</td></tr>
<tr><td>2</td><td>226-275</td><td>At this level, the medium of texts may be digital or printed, and texts may comprise continuous, non-continuous, or mixed types. Tasks at this level require respondents to make matches between the text and information, and may require paraphrasing or low-level inferences. Some competing pieces of information may be present. Some tasks require the respondent to
• cycle through or integrate two or more pieces of information based on criteria;
• compare and contrast or reason about information requested in the question; or
• navigate within digital texts to access-and-identify information from various parts of a document.</td><td>Tasks at this level require the respondent to identify and act on mathematical information and ideas embedded in a range of common contexts where the mathematical content is fairly explicit or visual with relatively few distractors. Tasks tend to require the application of two or more steps or processes involving calculation with whole numbers and common decimals, percents and fractions; simple measurement and spatial representation; estimation; and interpretation of relatively simple data and statistics in texts, tables and graphs.</td></tr>
<tr><td>3</td><td>276-325</td><td>Texts at this level are often dense or lengthy, and include continuous, non-continuous, mixed, or multiple pages of text. Understanding text and rhetorical structures become more central to</td><td>Tasks at this level require the respondent to understand mathematical information that may be less explicit, embedded in contexts that</td></tr>
</table>

		successfully completing tasks, especially navigating complex digital texts. Tasks require the respondent to identify, interpret, or evaluate one or more pieces of information, and often require varying levels of inference. Many tasks require the respondent to construct meaning across larger chunks of text or perform multi-step operations in order to identify and formulate responses. Often tasks also demand that the respondent disregard irrelevant or inappropriate content to answer accurately. Competing information is often present, but it is not more prominent than the correct information.	are not always familiar and represented in more complex ways. Tasks require several steps and may involve the choice of problem-solving strategies and relevant processes. Tasks tend to require the application of number sense and spatial sense; recognizing and working with mathematical relationships, patterns, and proportions expressed in verbal or numerical form; and interpretation and basic analysis of data and statistics in texts, tables and graphs.
4	326-375	Tasks at this level often require respondents to perform multiple-step operations to integrate, interpret, or synthesize information from complex or lengthy continuous, non-continuous, mixed, or multiple type texts. Complex inferences and application of background knowledge may be needed to perform the task successfully. Many tasks require identifying and understanding one or more specific, non-central idea(s) in the text in order to interpret or evaluate subtle evidence-claim or persuasive discourse relationships. Conditional information is frequently present in tasks at this level and must be taken into consideration by the respondent. Competing information is present and sometimes seemingly as prominent as correct information.	Tasks at this level require the respondent to understand a broad range of mathematical information that may be complex, abstract or embedded in unfamiliar contexts. These tasks involve undertaking multiple steps and choosing relevant problem-solving strategies and processes. Tasks tend to require analysis and more complex reasoning about quantities and data; statistics and chance; spatial relationships; and change, proportions and formulas. Tasks in this level may also require understanding arguments or communicating well-reasoned explanations for answers or choices.
5	376-500	At this level, tasks may require the respondent to search for and integrate information across multiple, dense texts; construct syntheses of similar and contrasting ideas or points of view; or evaluate evidence based arguments. Application and evaluation of logical and conceptual models of ideas may be required to accomplish tasks. Evaluating reliability of evidentiary sources and selecting key information is frequently a key requirement. Tasks often require respondents to be aware of subtle, rhetorical cues and to make high-level inferences or use specialized background knowledge.	Tasks at this level require the respondent to understand complex representations and abstract and formal mathematical and statistical ideas, possibly embedded in complex texts. Respondents may have to integrate multiple types of mathematical information where considerable translation or interpretation is required; draw inferences; develop or work with mathematical arguments or models; and justify, evaluate and critically reflect upon solutions or choices.

Tables 5.3 and 5.4 show the probability that adults with particular proficiency scores will complete items of different levels of difficulty in the domains of literacy and numeracy. For example, an adult with a proficiency score of 300 in literacy (i.e. the mid-point of level 3) has a 68% chance of successfully completing items of level 3 difficulty. He or she has a 29% chance of completing items of level 4 difficulty and a 90% probability of successfully completing items of level 2 difficulty.

Table 5.3 Probability of successfully completing items at different difficulty levels by proficiency score: Literacy

Item difficulty	Proficiency Score											
LEVEL	**150**	**175**	**200**	**225**	**250**	**275**	**300**	**325**	**350**	**375**	**400**	**425**
1	0.56	0.68	0.78	0.86	0.92	0.95	0.97	0.98	0.99	0.99	1.00	1.00
2	0.08	0.15	0.27	0.44	0.63	0.80	0.90	0.95	0.98	0.99	0.99	1.00
3	0.01	0.03	0.06	0.13	0.26	0.46	0.68	0.83	0.92	0.96	0.98	0.99
4	0.01	0.01	0.02	0.05	0.09	0.16	0.29	0.47	0.65	0.80	0.90	0.95

Table 5.4 Probability of successfully completing items at different difficulty levels by proficiency score: Numeracy

Item difficulty	Proficiency Score											
LEVEL	**150**	**175**	**200**	**225**	**250**	**275**	**300**	**325**	**350**	**375**	**400**	**425**
1	0.47	0.60	0.72	0.82	0.89	0.93	0.96	0.98	0.99	0.99	1.00	1.00
2	0.11	0.20	0.33	0.49	0.66	0.80	0.89	0.94	0.97	0.98	0.99	1.00
3	0.02	0.04	0.08	0.15	0.26	0.43	0.63	0.80	0.90	0.95	0.98	0.99
4	0.02	0.03	0.05	0.08	0.14	0.24	0.37	0.54	0.69	0.80	0.88	0.93

5.3.4 Problem solving in technology-rich environments

The problem-solving proficiency scale was divided into three levels. The problem solving in technology-rich environments framework (OECD 2010) identifies three main dimensions along which problems vary in quality and complexity. These are (1) the technology dimension, (2) the task dimension and (3) the cognitive dimension. Variations along each of these dimensions contribute to the overall difficulty of a problem. For instance, a problem is likely to be more complex if it involves the combined use of more than one computer application (e.g., e-mail and a spreadsheet); similarly, a problem is more complex if the task is defined in vague terms, as opposed to fully specified. Finally, a problem is likely to be more difficult if the respondent has to generate lots of deductions and inferences than if he or she just has to assemble or match different pieces of explicit information. The relationship between these dimensions and the proficiency levels is presented in Table 5.5. The descriptors of the levels are presented in Table 5.6.

Table 5.5 Technology, task and cognitive features of problems at each of the three main levels of proficiency

	Technology features	Task features	Cognitive processes
Level 1	• Generic applications • Little or no navigation required • Relevant information is directly available • Use of facilitating tools not required	• Few steps • Single operators	• Reach a given goal • Apply explicit criteria • Minimal monitoring demands • Simple relevance match • Categorical reasoning • No integration or transformation
Level 2	• Both generic and novel applications (e.g. web-based services) • Some navigation required to acquire information or perform actions • Use of tools facilitates operations	• Multiple steps • Multiple operators	• Goal may need to be defined • Apply explicit criteria • Generally higher monitoring demands • Generally involves resolving impasses • Some evaluation of relevance • Some integration or transformation • Inferential reasoning
Level 3	• Generic and novel applications • Some navigation required to acquire information or perform actions • Use of tools required to efficiently solve the problem	• Multiple steps • Multiple operators	• Goal may need to be defined • Establish and apply criteria • Generally high monitoring • High inferential reasoning and integration • Evaluate relevance and reliability • Generally involves resolving impasses

Table 5.6 Proficiency levels: Problem solving in technology-rich environments

Level	Score range	The types of tasks completed successfully at each level of proficiency
Below Level 1	0-240	Tasks are based on well-defined problems involving the use of only one function within a generic interface to meet one explicit criterion without any categorical, inferential reasoning or transforming of information. Few steps are required and no sub goal has to be generated.
1	241-290	At this level, tasks typically require the use of widely available and familiar technology applications, such as e-mail software or a web browser. There is little or no navigation required to access the information or commands required to solve the problem. The problem may be solved regardless of the respondent's awareness and use of specific tools and functions (e.g. a sort function). The tasks involve few steps and a minimal number of operators. At the cognitive level, the respondent can readily infer the goal from the task statement; problem resolution requires the respondent to apply explicit criteria; and there are few monitoring demands (e.g. the respondent does not have to check whether he or she has used the appropriate procedure or made progress towards the solution). Identifying contents and operators can be done through simple match. Only simple forms of reasoning, such as assigning items to categories, are required; there is no need to contrast or integrate information.
2	291-340	At this level, tasks typically require the use of both generic and more specific technology applications. For instance, respondents may have to make use of a novel online form. Some navigation across pages and applications is required to solve the problem. The use of tools (e.g. a sort function) can facilitate the resolution of the problem. The task may involve multiple steps and operators. The goal of the problem may have to be defined by the respondent, though the criteria to be met are explicit. There are higher monitoring demands. Some unexpected outcomes or impasses may appear. The task may require evaluating the relevance of a set of items to discard distractors. Some integration and inferential reasoning may be needed.
3	341-500	At this level, tasks typically require the use of both generic and more specific technology applications. Some navigation across pages and applications is required to solve the problem. The use of tools (e.g. a sort function) is required to make progress towards the solution. The task may involve multiple steps and operators. The goal of the problem may have to be defined by the respondent, and the criteria to be met may or may not be explicit. There are typically high monitoring demands. Unexpected outcomes and impasses are likely to occur. The task may require evaluating the relevance and reliability of information in order to discard distractors. Integration and inferential reasoning may be needed to a large extent.

Table 5.7 shows the probability of adults with particular proficiency in problem solving in technology-rich environments completing problem solving items of different levels of difficulty.

Table 5.7 Probability of successfully completing items at different difficulty levels by proficiency score: Problem solving in technology-rich environments

Item Difficulty	**Proficiency Score**									
LEVEL	**190**	**215**	**240**	**265**	**290**	**315**	**340**	**365**	**390**	**415**
1	0.02	0.06	0.17	0.40	0.69	0.87	0.95	0.98	0.99	1.00
2	0.03	0.05	0.10	0.19	0.35	0.56	0.76	0.88	0.94	0.97
3	0.00	0.01	0.02	0.05	0.13	0.29	0.49	0.67	0.80	0.87

5.3.5 A note about the reporting of problem solving in technology-rich environments

The populations for whom proficiency scores for problem solving in technology-rich environments are reported *are not identical* across countries. Proficiency scores relate only to the proportion of the target population in each participating country that was able to undertake the computer-based version of the assessment, and thus meets the preconditions for displaying competence in this domain.

Four groups of respondents did not take the computer-based assessment[4], those who:

- indicated in completing the background questionnaire that they had never used a computer (group 1);
- had some experience with computers but who "failed" the ICT core assessment designed to determine whether a respondents had the basic computer skills necessary to undertake the computer-based assessment (group 2);
- had some experience with computers but opted not to take the computer-based assessment (group 3); or
- did not attempt the ICT core for literacy-related reasons (group 4).

By definition, a minimum level of competence in the use of computer tools and applications and a minimum level of proficiency in literacy and numeracy is required in order to display proficiency in problem solving in technology-rich environments. Individuals in groups 1 and 2 are, thus, treated as not meeting the necessary preconditions for displaying proficiency and have no proficiency score in the domain of problem solving.

Respondents who did not attempt the ICT core for literacy-related reasons (group 4) have not been attributed a problem-solving score due to lack of sufficient information.

Respondents who opted not to take the computer-based assessment (group 3), however, represent a different category. They are individuals who, on their own initiative, decided to take the paper-and-pencil version of the assessment without going through the process designed to direct respondents to the

[4] Defined as taking, at a minimum, the core literacy and numeracy assessments on the computer.

computer-based or paper-and-pencil pathways of the assessment. As a result, it is not known whether or not they possessed the computer skills necessary to complete the computer-based assessment.

Three options for how to treat this group were considered: imputing their proficiency in problem solving on the basis of their proficiency in literacy and numeracy and their background characteristics; treating them as non-respondents; or reporting them as a separate category of the group that could not display competence. The latter option was adopted. Imputation was rejected on the grounds that refusals appeared to have different characteristics to respondents taking the computer-based assessment pathway. In fact, they appeared to be more similar to the respondents who did not have computer skills than to those who took the computer-based assessment. The option of treating them as non-respondents was rejected for similar reasons.

In reporting the results concerning problem solving in technology-rich environments, the following approach was adopted:

- When reporting proficiency in problem solving in technology-rich environments on the continuous scale at the country level, the proportion of the population displaying proficiency is reported in conjunction with country-level statistics (e.g. means, standard deviations, etc).
- When reporting distributions of the population by proficiency levels, information is presented for the entire adult population as a whole (i.e. those displaying proficiency plus those not displaying proficiency). The number or proportion of the population not displaying proficiency is always reported when results are presented by proficiency level.

5.4 An overview of literacy, numeracy and problem solving in technology-rich environments

Groups of experts in their fields developed the frameworks for each of the skills domains assessed in the PIAAC. They guided the development and selection of assessment items and the interpretation of results. Their work is presented in *Literacy, Numeracy and Problem Solving in Technology-Rich Environments: Framework for the OECD Survey of Adult Skills* (OECD 2012).[5] The frameworks define and describe what is measured. In each case, three main dimensions are identified:

- content – the artifacts, tools, knowledge, representations, cognitive challenges, etc. that constitute the corpus adults must respond to or use;
- cognitive strategies – the processes that adults must bring into play to respond to or use given content in an appropriate manner; and
- context – the different situations in which adults have to read, display numerate behaviour, and solve problems.

Table 5.8 provides an overview of each of the three domains, including a definition and the content, cognitive strategies and context related to each. These are described in more detail in the remainder of this chapter.

[5] For the complete framework documents, see PIAAC Literacy Expert Group (2009), PIAAC Numeracy Expert Group (2009), PIAAC Expert Group in Problem Solving in Technology-Rich Environments (2009), and Sabatini and Bruce (2009).

Table 5.8 Summary of assessment domains in the PIAAC

	Literacy	**Numeracy**	**Problem solving in technology-rich environments**
Definition	Literacy is defined as the ability to understand, evaluate, use and engage with *written texts* to participate in society, to achieve one's goals, and to develop one's knowledge and potential. Literacy encompasses a range of skills from the decoding of written words and sentences to the comprehension, interpretation, and evaluation of complex texts. It does not, however, involve the production of text (writing[i]). Information on the skills of adults with low levels of proficiency is provided by an assessment of reading components that covers text vocabulary, sentence comprehension and passage fluency.	Numeracy is defined as the ability to access, use, interpret and communicate mathematical information and ideas in order to engage in and manage the mathematical demands of a range of situations in adult life. To this end, numeracy involves managing a situation or solving a problem in a real context, by responding to mathematical content/information/ideas represented in multiple ways.	Problem solving in technology-rich environments is defined as the ability to use digital technology, communication tools and networks to acquire and evaluate information, communicate with others and perform practical tasks. The assessment focuses on the abilities to solve problems for personal, work and civic purposes by setting up appropriate goals and plans, and accessing and making use of information through computers and computer networks.
Content	Different types of text. Texts are characterized by their medium (print-based or digital) and by their format: • Continuous or prose texts • Non-continuous or document texts • Mixed texts • Multiple texts	Mathematical content, information and ideas: • Quantity and number • Dimension and shape • Pattern, relationships and change • Data and chance Representations of mathematical information: • Objects and pictures • Numbers and symbols • Visual displays (e.g. diagrams, maps, graphs, tables) • Texts • Technology-based displays	Technology: • Hardware devices • Software applications • Commands and functions • Representations (e.g. text, graphics, video) Tasks: • Intrinsic complexity • Explicitness of the problem statement
Cognitive strategies	• Access and identify • Integrate and interpret (relating parts of text to one another) • Evaluate and reflect	• Identify, locate or access • Act upon and use (order, count, estimate, compute, measure, model) • Interpret, evaluate and analyze • Communicate	• Set goals and monitoring progress • Plan • Acquire and evaluate information • Use information

Contexts	• Work-related • Personal • Society and community • Education and training	• Work-related • Personal • Society and community • Education and training	• Work-related • Personal • Society and community • Education and training

5.4.1 Literacy

5.4.1.1 Definition

In the PIAAC, literacy is defined as "understanding, evaluating, using and engaging with written texts to participate in society, to achieve one's goals, and to develop one's knowledge and potential".

Key to this definition is the fact that literacy is defined in terms of the *reading* of *written text*s and does not involve either the comprehension or production of spoken language or the production of text (writing). While literacy is commonly seen as encompassing the ability to write as well as read (see UNESCO 2005), the dimension of writing is not part of the construct measured in the OECD survey. This is largely because of the difficulty of assessing writing in a reliable and valid way in an international comparative assessment. In addition, literacy is conceived as a skill that involves constructing meaning, and evaluating and using texts to achieve a range of possible goals in a variety of contexts. In other words, in the PIAAC, literacy extends well beyond the skills of decoding or comprehending texts to using them appropriately in context.

5.4.1.2 Content

The corpus of texts to which adults are required to respond are classified along two principle axes: medium and format. Medium refers to the nature of the support in which a text is instantiated or displayed. Format refers to the organizational and structural features of texts, whether digital or print-based.

In terms of medium, texts are classified as either digital or print-based. Digital texts are texts that are stored as digital information (a series of 1s and 0s) and accessed in the form of screen-based displays on devices such as computers and smart phones. Print-based texts are texts printed on paper or other material supports; these include newspapers, books, pamphlets and road signs. Digital texts have a range of features, in addition to being displayed on screens, that distinguishes them from print-based texts. These include hypertext links to other documents, specific navigation features (e.g. scroll bars, use of menus) and interactivity. The PIAAC is the first international assessment of adult skills to incorporate the reading of digital texts as part of the construct of (reading) literacy.

In terms of *format*, texts are categorized in the following way:

- continuous texts, which are made up of sentences organized in paragraphs that incorporate a range of rhetorical stances, such as description, narration, instruction and argumentation;
- non-continuous texts, which are organized in a matrix format or around graphic features. Several different organizing structures are identified, including simple and complex lists, graphic documents (e.g. graphs, diagrams), locative documents (e.g. maps) and entry documents (e.g. forms);

- mixed texts, which involve combinations of continuous and non-continuous elements (e.g. a newspaper article or a webpage that includes text and graphics); and
- multiple texts, which consist of juxtaposing or linking independently generated elements, such as an e-mail that contains a record of the separate messages that constitute an exchange over a period of time, or a blog post that contains an initial text and a string of related texts consisting of comments in response to the initial text and comments in response to other comments.

5.4.1.3 Cognitive strategies

Readers generally use three broad cognitive strategies when responding to written texts:

- access and identify;
- integrate and interpret; and
- evaluate and reflect.

Accessing and identifying involves locating information in a text. At one extreme, this can be a relatively simple operation when the information sought is clearly identified. At the other, it can be a complicated operation requiring inferential reasoning and an understanding of rhetorical strategies.

Integrating and interpreting involves understanding the relationships between different parts of a text to construct meaning and draw inferences from the text as a whole.

Evaluating and reflecting requires the reader to relate the information in the text to other information, knowledge and experiences, for example, to assess the relevance or credibility of a text.

5.4.1.4 Contexts

Adults read materials in a variety of contexts that affect the types of texts they encounter, the nature of the content, the motivation to read, and the manner in which texts are interpreted. The texts selected for the literacy assessment are related to four broad contexts:

- work-related;
- personal;
- community and society; and
- education and training.

Texts related to *work and occupation* include materials that discuss job search, wages, salaries and other benefits, and the experience of work.

Materials in the area of *personal* uses include texts concerning the home and family (e.g. interpersonal relationships, personal finances, housing and insurance); health and safety (e.g. drugs and alcohol, disease prevention and treatment, safety and accident prevention, first aid, emergencies, and lifestyle); consumer economics (e.g. banking, savings, advertising, prices); and leisure and recreation (e.g. travel, recreational activities).

Texts related to *community and society* includes materials that deal with public services, government, community groups and activities, and current events. Materials related to *education and training* cover text which refer to learning opportunities for adults or others.

5.4.1.5 Distribution of test items by task characteristics

Tables 5.9-5.11 below show the distribution of the literacy assessment items in the PIAAC by task characteristics. The final selection of items was determined taking into account the following factors: the performance of items in the field test, the need to cover the main dimensions of literacy as defined by the assessment frameworks, the need to include sufficient items that had been used in previous surveys to ensure comparability of the results, and the constraints imposed by the assessment design.

Table 5.9 Distribution of literacy items by medium

	Final item set	
	No.	**%**
Print-based texts	30	52
Digital texts	28	48
Total	58	100

Note: Each category includes continuous, non-continuous and combined texts.

Table 5.10 Distribution of literacy items by context

	Final item set	
	No.	**%**
Work	10	17
Personal	29	50
Community	10	23
Education	6	10
Total	58	100

Table 5.11 Distribution of literacy items by cognitive strategy

	Final item set	
	No.	**%**
Access and identify	31	53
Integrate and interpret	18	31
Evaluate and reflect	9	16
Total	58	100

Literacy sample items

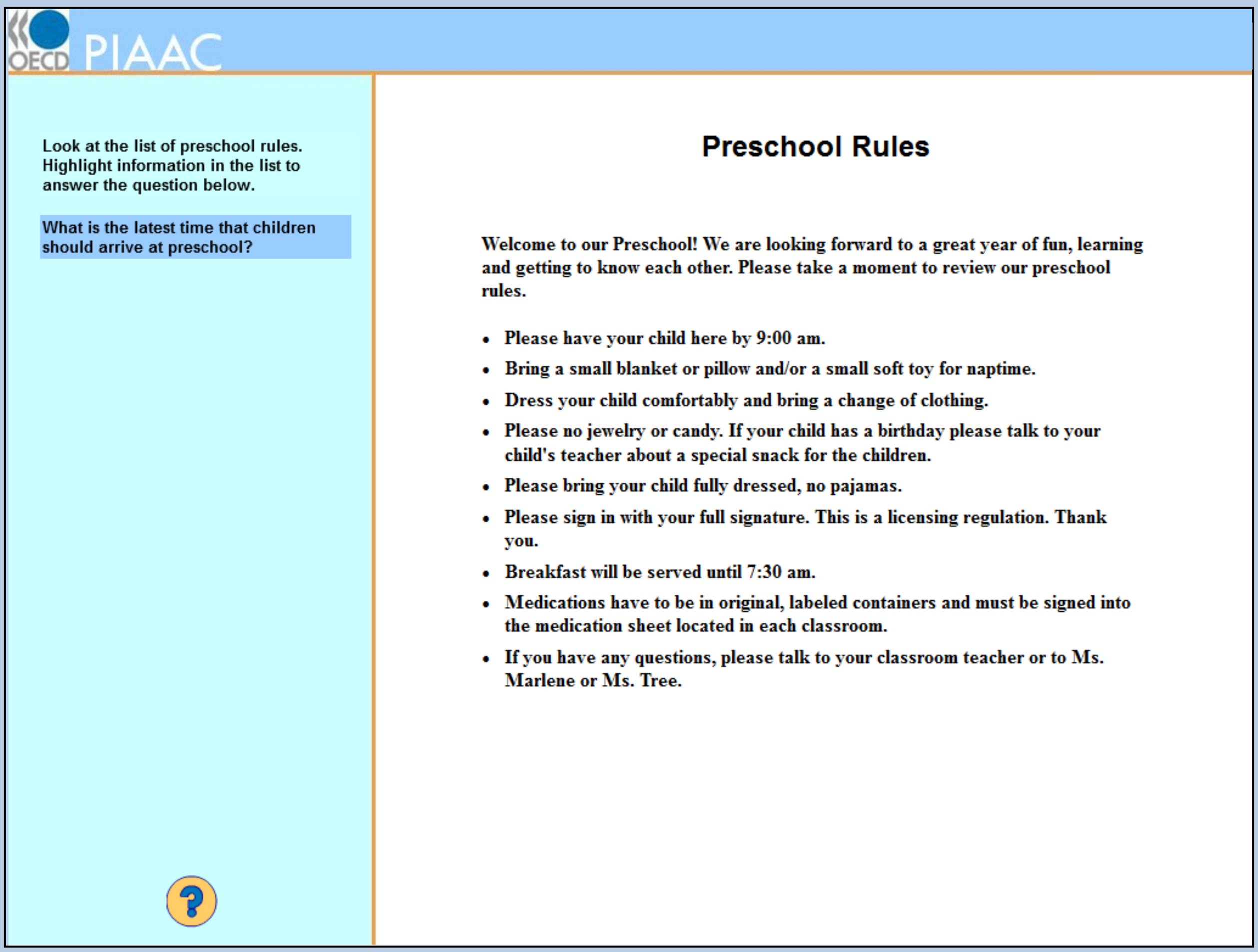
Two examples of the literacy items used in the PIAAC are presented below. Both use print-based stimuli. The sample problem-solving items presented further below give an idea of the type of "digital" stimulus material used.

The items are presented in the form delivered by the computer-based version of the assessment. To answer the questions, respondents highlighted words and phrases or clicked on the appropriate location on the screen using a mouse.

Sample Item 1: Preschool Rules

"Preschool Rules" represents an easy item and focuses on the following aspects of the literacy construct:

Medium	*Print*
Context	*Personal*
Cognitive strategy	*Access and identify*



Look at the list of preschool rules. Highlight information in the list to answer the question below.

What is the latest time that children should arrive at preschool?

Preschool Rules

Welcome to our Preschool! We are looking forward to a great year of fun, learning and getting to know each other. Please take a moment to review our preschool rules.

- **Please have your child here by 9:00 am.**
- **Bring a small blanket or pillow and/or a small soft toy for naptime.**
- **Dress your child comfortably and bring a change of clothing.**
- **Please no jewelry or candy. If your child has a birthday please talk to your child's teacher about a special snack for the children.**
- **Please bring your child fully dressed, no pajamas.**
- **Please sign in with your full signature. This is a licensing regulation. Thank you.**
- **Breakfast will be served until 7:30 am.**
- **Medications have to be in original, labeled containers and must be signed into the medication sheet located in each classroom.**
- **If you have any questions, please talk to your classroom teacher or to Ms. Marlene or Ms. Tree.**

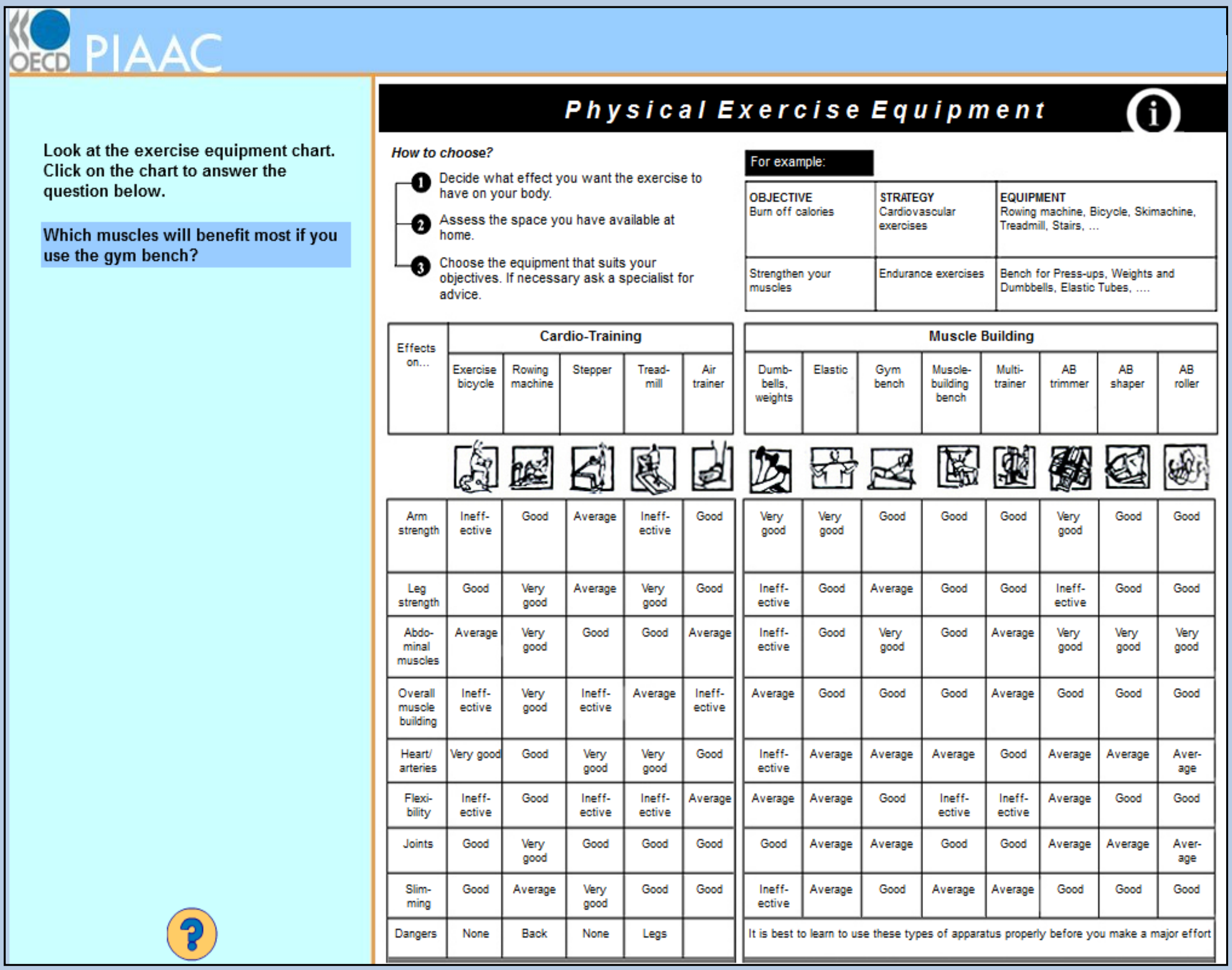
Sample Items 2 and 3: Physical Exercise Equipment

In many cases, several questions are associated with the same stimulus material. In the case of the stimulus relating to physical exercise equipment, there are two associated questions or test items.

The first item represents a relatively easy item and focuses on the following aspects of the literacy construct:

Medium	*Print*
Context	*Personal*
Cognitive strategy	*Access and identify*

Respondents answer the question by clicking on the cell in the chart that contains information about exercise equipment. Each of the cells and all of the images can be highlighted by clicking on them and multiple cells can be selected.



PIAAC

Look at the exercise equipment chart. Click on the chart to answer the question below.

Which muscles will benefit most if you use the gym bench?

Physical Exercise Equipment

How to choose?

1. Decide what effect you want the exercise to have on your body.
2. Assess the space you have available at home.
3. Choose the equipment that suits your objectives. If necessary ask a specialist for advice.

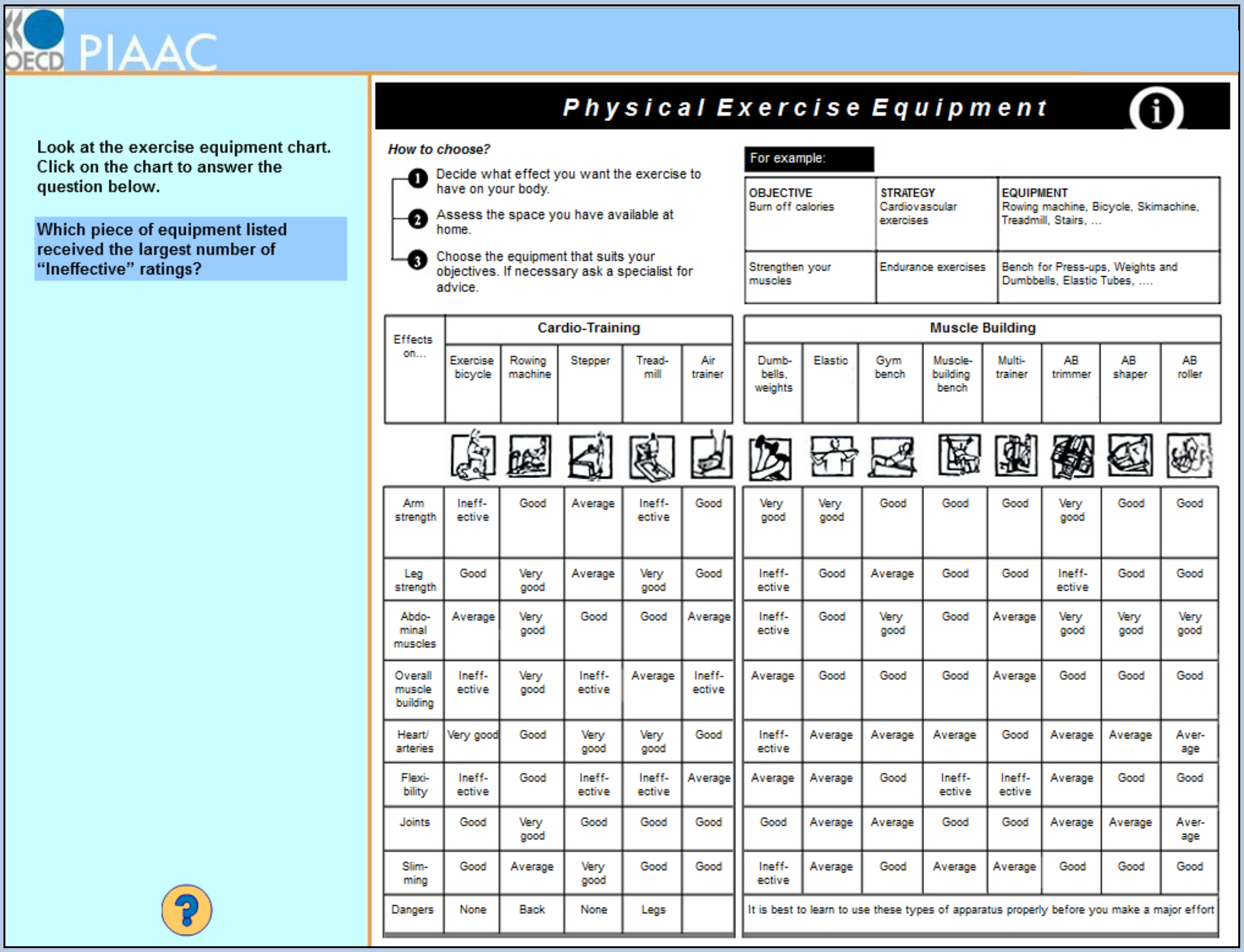
For example:

OBJECTIVE	STRATEGY	EQUIPMENT
Burn off calories	Cardiovascular exercises	Rowing machine, Bicycle, Skimachine, Treadmill, Stairs, ...
Strengthen your muscles	Endurance exercises	Bench for Press-ups, Weights and Dumbbells, Elastic Tubes, ...

Effects on...	Cardio-Training					Muscle Building							
	Exercise bicycle	Rowing machine	Stepper	Tread-mill	Air trainer	Dumb-bells, weights	Elastic	Gym bench	Muscle-building bench	Multi-trainer	AB trimmer	AB shaper	AB roller
Arm strength	Ineff-ective	Good	Average	Ineff-ective	Good	Very good	Very good	Good	Good	Good	Very good	Good	Good
Leg strength	Good	Very good	Average	Very good	Good	Ineff-ective	Good	Average	Good	Good	Ineff-ective	Good	Good
Abdo-minal muscles	Average	Very good	Good	Good	Average	Ineff-ective	Good	Very good	Good	Average	Very good	Very good	Very good
Overall muscle building	Ineff-ective	Very good	Ineff-ective	Average	Ineff-ective	Average	Good	Good	Good	Average	Good	Good	Good
Heart/ arteries	Very good	Good	Very good	Very good	Good	Ineff-ective	Average	Average	Average	Good	Average	Average	Aver-age
Flexi-bility	Ineff-ective	Good	Ineff-ective	Ineff-ective	Average	Average	Average	Good	Ineff-ective	Ineff-ective	Average	Good	Good
Joints	Good	Very good	Good	Good	Good	Good	Average	Average	Good	Good	Average	Average	Aver-age
Slim-ming	Good	Average	Very good	Good	Good	Ineff-ective	Average	Good	Average	Average	Good	Good	Good
Dangers	None	Back	None	Legs		It is best to learn to use these types of apparatus properly before you make a major effort							

The second item represents a relatively easy item and focuses on the following aspects of the literacy construct:

Medium	*Print*
Context	*Personal*
Cognitive strategy	*Integrate and interpret*



PIAAC

Look at the exercise equipment chart. Click on the chart to answer the question below.

Which piece of equipment listed received the largest number of "Ineffective" ratings?

Physical Exercise Equipment

How to choose?

1. Decide what effect you want the exercise to have on your body.
2. Assess the space you have available at home.
3. Choose the equipment that suits your objectives. If necessary ask a specialist for advice.

For example:

OBJECTIVE	STRATEGY	EQUIPMENT
Burn off calories	Cardiovascular exercises	Rowing machine, Bicycle, Skimachine, Treadmill, Stairs, …
Strengthen your muscles	Endurance exercises	Bench for Press-ups, Weights and Dumbbells, Elastic Tubes, ….

Effects on…	Cardio-Training					Muscle Building							
	Exercise bicycle	Rowing machine	Stepper	Tread-mill	Air trainer	Dumb-bells, weights	Elastic	Gym bench	Muscle-building bench	Multi-trainer	AB trimmer	AB shaper	AB roller
Arm strength	Ineff-ective	Good	Average	Ineff-ective	Good	Very good	Very good	Good	Good	Good	Very good	Good	Good
Leg strength	Good	Very good	Average	Very good	Good	Ineff-ective	Good	Average	Good	Good	Ineff-ective	Good	Good
Abdo-minal muscles	Average	Very good	Good	Good	Average	Ineff-ective	Good	Very good	Good	Average	Very good	Very good	Very good
Overall muscle building	Ineff-ective	Very good	Ineff-ective	Average	Ineff-ective	Average	Good	Good	Good	Average	Good	Good	Good
Heart/ arteries	Very good	Good	Very good	Very good	Good	Ineff-ective	Average	Average	Average	Good	Average	Average	Aver-age
Flexi-bility	Ineff-ective	Good	Ineff-ective	Ineff-ective	Average	Average	Average	Good	Ineff-ective	Ineff-ective	Average	Good	Good
Joints	Good	Very good	Good	Good	Good	Good	Average	Average	Good	Good	Average	Average	Aver-age
Slim-ming	Good	Average	Very good	Good	Good	Ineff-ective	Average	Good	Average	Average	Good	Good	Good
Dangers	None	Back	None	Legs		It is best to learn to use these types of apparatus properly before you make a major effort							

5.4.2 Reading Components

To provide more detailed information about adults with poor literacy skills, the survey's literacy assessment is complemented by a test of "reading component" skills. Reading components are the basic set of decoding skills that are essential for extracting meaning from written texts: knowledge of vocabulary (word recognition), the ability to process meaning at the level of the sentence, and fluency in reading passages of text. Skilled readers are able to undertake these types of operations automatically. To assess this skill, the time taken by respondents to complete the tasks was recorded.

Examples of reading component items

Print vocabulary

Items testing print vocabulary consist of a picture of an object and four printed words, one of which refers to the pictured object. Respondents are asked to circle the word that matches the picture.

ear	egg	lip	jar

Sentence processing

The sentence-processing items require the respondent to assess whether a sentence makes sense in terms of the properties of the real world or the internal logic of the sentence. The respondent reads the sentence and circles YES if the sentence makes sense or NO if the sentence does not make sense.

Three girls ate the song.	**YES**	**NO**
The man drove the green car.	**YES**	**NO**

The lightest balloon floated in the bright sky.	**YES**	**NO**
A comfortable pillow is soft and rocky.	**YES**	**NO**
A person who is twenty years old is older than a person who is thirty years old.	**YES**	**NO**

Passage comprehension

In items assessing passage comprehension, respondents are asked to read a passage in which they are required, at certain points, to select the word from the two alternatives provided that makes the most sense.

To the editor: Yesterday, it was announced that the cost of riding the bus will increase. The price will go up by twenty percent starting next wife / month. As someone who rides the bus every day, I am upset by this foot / increase. I understand that the cost of gasoline / student has risen. I also understand that riders have to pay a fair price / snake for bus service. I am willing to pay a little more because I rely on the bus to get to object / work. But an increase / uncle of twenty percent is too much.

This increase is especially difficult to accept when you see the city's plans to build a new sports stadium. The government will spend millions on this project even though we already have a science / stadium. If we delay the stadium, some of that money can be used to offset the increase in bus fares / views. Then, in a few years, we can decide if we really do need a new sports cloth / arena. Please let the city council know you care about this issue by attending the next public meeting / frames.

5.4.3 Numeracy

The PIAAC defines numeracy as “the ability to access, use, interpret and communicate mathematical information and ideas, in order to engage in and manage the mathematical demands of a range of situations in adult life” (OECD 2012). Numeracy is further defined in terms of the concept of “numerate behaviour” that involves managing a situation or solving a problem in a real context by responding to mathematical information and content represented in various ways.

It is recognized that literacy skills such as reading and writing enable numerate behaviour, and that when mathematical representations involve text, performance on numeracy tasks is, in part, dependent on the ability to read and understand text. However, numeracy in the PIAAC involves more than applying arithmetical skills to information embedded in text. In particular, numeracy relates to a wide range of skills and knowledge, not just arithmetic knowledge and computation, a range of responses that may involve more than numbers, and responses to a range of representations, not just numbers in texts.

5.4.3.1 Content

The survey covers the following areas of mathematical content, information and ideas:

- quantity and number;
- dimension and shape;
- pattern, relationships and change; and
- data and chance.

Quantity encompasses attributes such as the number of features or items, prices, size (e.g. length, area and volume), temperature, humidity, atmospheric pressure, populations and growth rates, revenues and profit, etc. *Number* is fundamental to quantification. Numbers (whether whole numbers or fractions, decimals or percents) serve as counters or estimators, indicate parts or comparisons. Positive and negative numbers can also serve as directional indicators. In calculations, operations (i.e. the four main operations of +, –, x, ÷ and others, such as squaring) are performed on quantities and numbers.

Dimension covers the description of "things" in space, such as projections, lengths, perimeters, areas, planes, surfaces, location, etc. *Shape* involves a category describing real images and entities that can be visualized in two or three dimensions (e.g. houses and buildings, designs in art and craft, safety signs, packaging, snowflakes, knots, crystals, shadows and plants).

Pattern covers regularities encountered in the world, such as those in musical forms, nature, traffic, etc. *Relationships* and *change* relate to the mathematics of how things in the world are associated with one another or develop over time.

Data and chance encompass two separate but related topics. *Data* covers the "big ideas" related to variability, sampling, error, prediction and statistical topics, such as data collection, data displays and graphs. *Chance* covers the "big ideas" related to probability and relevant statistical methods.

5.4.3.2 Representations of mathematical information

In the PIAAC, mathematical information may be represented in the form of:

- objects and pictures;
- numbers and symbols;
- visual displays; texts; and
- technology-based displays.

Objects (physical entities) can be counted and measured. *Pictures* (e.g. photographs, paintings, videos) also represent mathematical information such as number, size, volume or location. *Numbers and symbols* include numerals, letters, and operation or relationship signs and formulae. *Visual displays* cover graphic presentations of mathematical information, such as diagrams or charts, graphs and tables (used to display aggregate statistical or quantitative information through objects, counting data, etc.) or maps (e.g. of a city or a project plan). Two different kinds of *text* may be encountered in numeracy tasks. The first involves representing mathematical information in textual form, i.e. as words or phrases that carry mathematical meaning. The second involves expressing mathematical information in mathematical notations or symbols (e.g. numbers, plus or minus signs, symbols for units of measure, etc.) that are surrounded by text that provides additional information and context.

5.4.3.3 Cognitive strategies

The following processes define the dimension of cognitive strategies:

- identify, locate, or access;
- act upon or use;
- interpret;
- evaluate/analyze; and
- communicate.

In virtually all situations, people have to *identify, locate or access* some mathematical information relevant to their purpose or goal. In isolation, this response type often requires a low level of mathematical understanding or the application of simple arithmetic skills. However, this response type is usually combined with the other types of responses listed below.

Acting upon or using involves the use of known mathematical procedures and rules, such as counting and making calculations. It may also call for ordering or sorting, estimating or using various measuring devices, or for using (or developing) a formula that serves as a model of a situation or a process.

Interpretation involves evaluating the meaning and implications of mathematical or statistical information (e.g. a graph showing variation in an exchange rate) and developing an opinion about the information.

Evaluation/analysis is in part an extension of interpretation. It involves analyzing a problem, evaluating the quality of the solution against some criteria or contextual demands and, if necessary, reviewing the interpretation, analysis and evaluation stages.

While defined as a cognitive process forming part of this dimension of the numeracy framework, the ability to *communicate* numerical and mathematical content is not assessed in the PIAAC.

5.4.3.4 Contexts

The items selected for the numeracy assessment are related to four contexts:

- work-related;
- personal;
- society or community; and
- education and training.

Representative tasks related to *work situations* include: completing purchase orders; totaling receipts; calculating change; managing schedules, budgets and project resources; using spreadsheets; organizing and packing goods of different shapes; completing and interpreting control charts; making and recording measurements; reading blueprints; tracking expenditures; predicting costs; and applying formulas.

Representative tasks related to the context of *personal life* include: handling money and budgets; shopping and managing personal time; planning travel; playing games of chance; understanding sports scoring and statistics; reading maps; and using measurements in home situations, such as cooking, doing home repairs or pursuing hobbies.

Adults need to have an awareness of what is occurring in the *society, the economy and the environment* (e.g. trends in crime, health, wages, pollution), and may have to take part in social events or community action. This requires a capacity to read and interpret quantitative information presented in the media, including statistical messages and graphs. Adults also have to manage a variety of situations, such as raising funds for a football club or interpreting the results of a study on a medical condition.

Competence in numeracy may enable a person to participate in *education and training*, whether for academic purposes or as part of vocational training. In either case, it is important to know some of the more formal aspects of mathematics that involve symbols, rules and formulae and to understand some of the conventions used to apply mathematical rules and principles.

5.4.3.5 Distribution of test items by task characteristics

Tables 5.12-5.14 below show the distribution of the numeracy assessment items included in the PIAAC by task characteristics. As in the case of literacy, the final selection of items reflected the performance of items in the field test, the need to cover the main dimensions of numeracy as defined by the assessment frameworks, the need to include sufficient items that had been used in previous surveys to ensure comparability of the results, and the constraints imposed by the assessment design.

Table 5.12 Distribution of numeracy items by response type

	Final item set	
	No.	**%**
Identify, locate or access	3	5
Act upon, use	34	61
Interpret, evaluate/analyze	19	34
Total	56	100

Table 5.13 Distribution of numeracy items by context

	Final item set	
	No.	**%**
Work-related	13	23
Personal	25	45
Society or community	14	25
Education and training	4	7
Total	56	100

Table 5.14 Distribution of numeracy items by mathematical content

	Final item set	
	No.	**%**
Data and chance	12	21
Dimension and shape	16	29
Pattern, relationships and change	15	27
Quantity and number	13	23
Total	56	100

Numeracy sample items

Sample Item 1: Births in the US

The items are presented in the form delivered by the computer-based version of the assessment. To answer the questions, respondents clicked in the appropriate box or typed figures in the spaces provided.

This item (of medium difficulty) focuses on the following aspects of the numeracy construct:

Content	*Data and chance*
Process	*Interpret, evaluate*
Context	*Community and society*

Respondents were asked to respond by clicking on one or more of the time periods provided in the left pane on the screen.

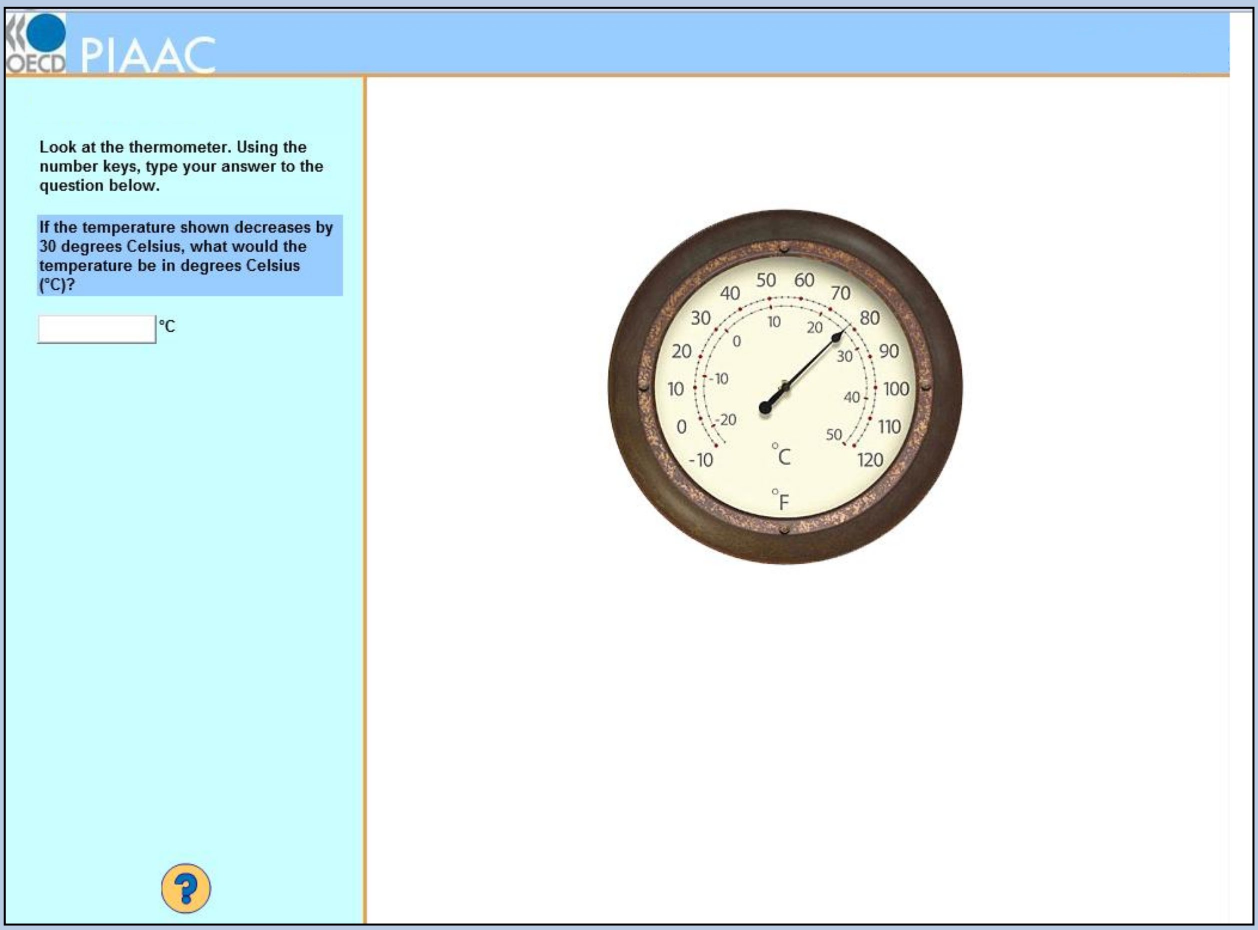
Correct Response: 1957 - 1967 and 1967 – 1977

Sample Item 2: Thermometer

This item (of low difficulty) focuses on the following aspects of the numeracy construct:

Content	*Dimension and shape*
Process	*Act upon, use (measure)*
Context	*Every day or work*

Respondents were asked to type in a numerical response based on the graphic provided.

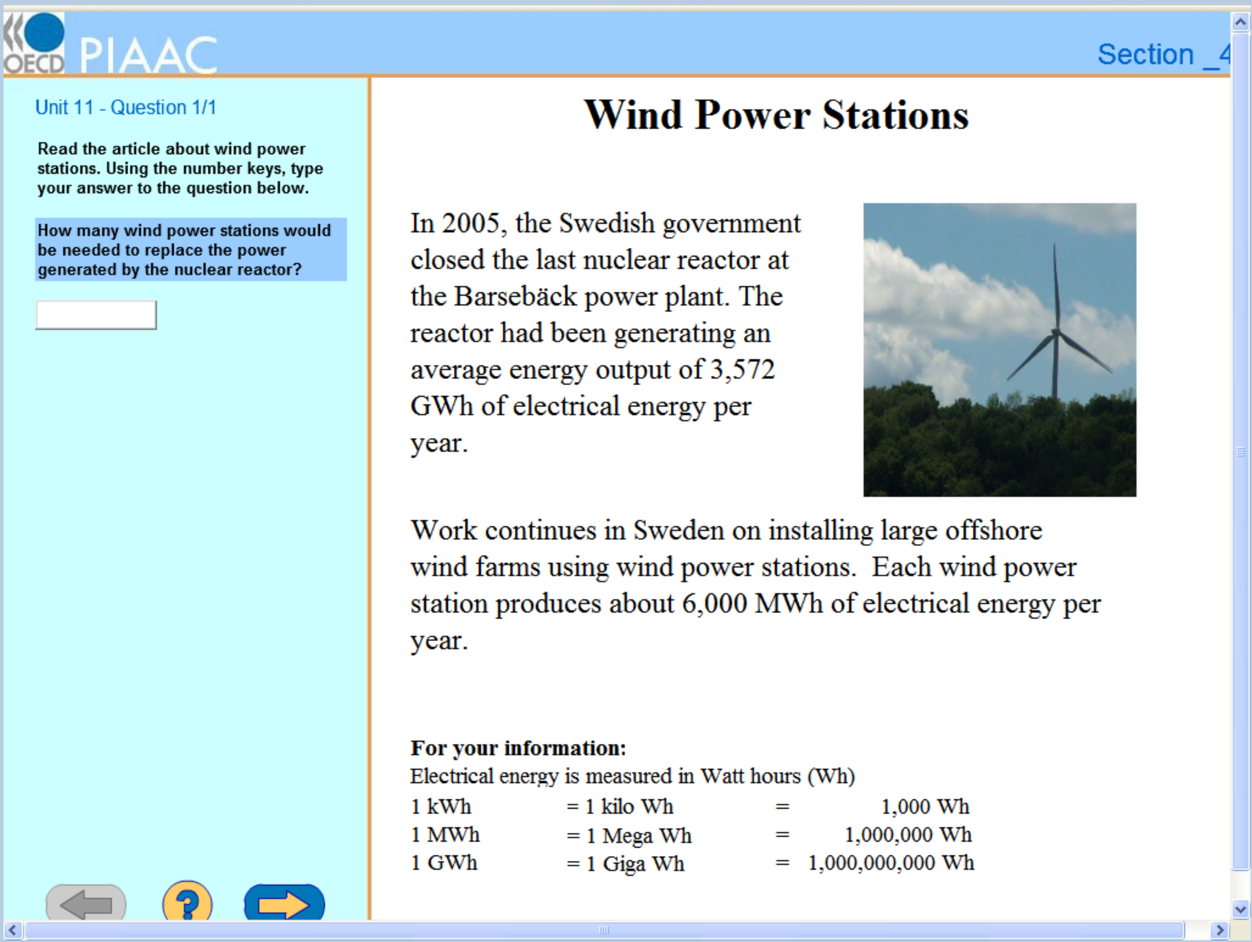


Correct Response: Any value between -4 and -5

Sample Item 3: Wind Power Stations

This sample item (of medium difficulty) focuses on the following aspects of the numeracy construct:

Content	*Quantity and Number*
Process	*Act upon, use (compute)*
Context	*Community and society*



5.4.4 Problem solving in technology-rich environments

In the PIAAC, problem solving in technology-rich environments is defined as "using digital technology, communication tools and networks to acquire and evaluate information, communicate with others and perform practical tasks". The first wave of the PIAAC focuses on "the abilities to solve problems for personal, work and civic purposes by setting up appropriate goals and plans, and accessing and making use of information through computers and computer networks" (OECD 2012).

The problem solving in technology-rich environments domain covers the specific types of problems people deal with when using ICT. These problems share the following characteristics:

- the problem is primarily a consequence of the availability of new technologies;
- the solution to the problem requires the use of computer-based artifacts (applications, representational formats, computational procedures); and
- the problems are related to technology-rich environments themselves (e.g. how to operate a computer, how to fix a settings problem, how to use an Internet browser).

Problem solving in technology-rich environments is a domain of competence that represents the intersection of what are sometimes described as "computer literacy" skills (i.e. the capacity to use ICT tools and applications) and the cognitive skills required to solve problems. Some basic knowledge regarding the use of ICT input devices (e.g. use of a keyboard and mouse and screen displays), file management tools, applications (word processing, e-mail), and graphic interfaces is essential for performing assessment tasks. However, the objective is not to test the use of ICT tools and applications in isolation, but rather to assess the capacity of adults to use these tools to access, process, evaluate and analyze information effectively.

5.4.4.1 Content

The content of the assessment encompasses two areas: technology and tasks.

Technology refers to the devices, applications and functionalities through which problem solving is conducted. It encompasses digital devices such as computers, mobile phones, GPS devices, software applications and the commands, functions and representations of information on which these applications depend. In the first cycle of the survey, only laptop computers with a limited number of simulated software applications – including e-mail, word processing, spreadsheets and websites – were used. For operational reasons, sound, animations and videos were not used.

Tasks are the circumstances that trigger a person's awareness and understanding of the problem and determine the actions needed to be taken in order to solve the problem. Ordinarily, a wide range of conditions can initiate problem solving. Tasks are defined in terms of intrinsic complexity and the explicitness of the problem statement. The *intrinsic complexity* of a problem is determined by:

- the minimum number of steps required to solve the problem;
- the number of options or alternatives at various stages in the solution path;
- the diversity of operators required to be used, and the complexity of computation/transformation;
- the likelihood of impasses or unexpected outcomes;

- the number of requirements that have to be satisfied to arrive at a solution; and
- the amount of transformation required to communicate a solution.

The *explicitness of the problem statement* relates to the extent to which the problem is ill-defined (the task is implicit and its components are largely unspecified) or well-defined (the task is explicit and its components are described in detail).

5.4.4.2 Cognitive strategies

The process aspect of the assessment relates to the mental structures and processes involved when a person solves a problem. These include setting goals and monitoring progress; planning; locating, selecting and evaluating information; and organizing and transforming information.

Setting goals and monitoring progress involves identifying objectives in the context of the constraints (explicit and implicit) of a situation; establishing and applying criteria for respecting constraints and arriving at a solution; monitoring progress; and detecting and interpreting unexpected events, impasses and breakdowns as one proceeds along the path to a solution.

Planning and self-organization covers the processes of setting up adequate plans, procedures and strategies (operators) and selecting appropriate devices, tools or categories of information.

Acquiring and evaluating information involves orienting and focusing attention, selecting information, assessing the reliability, relevance, adequacy and comprehensibility of information, and reasoning about sources and contents.

Using information involves organizing information; integrating information drawn from different and possibly inconsistent texts and from different formats; making informed decisions; transforming information through rewriting, from text to table, from table to graph, etc.; and communicating with relevant parties.

5.4.4.3 Contexts

The contexts are those of personal life, work and occupation, and civic participation.

5.4.4.4 Distribution of test items by task characteristics

Tables 5.15-5.20 below show the distribution of the problem-solving assessment items included in the PIAAC by task characteristics.

In total 14 items were administered in the assessment of problem solving in technology environments. Items consisted of scenarios that involved a number of sub-tasks such as searching through simulated web sites for relevant information or transferring information from one application to another. The time taken to complete the problem solving tasks was considerably longer than that in either literacy or numeracy.

Table 5.15 Distribution of problem-solving tasks by cognitive dimensions

Dimension	No.*
Setting goals and monitoring progress	4
Planning	7
Acquiring and evaluating information	8
Using information	6

*Does not add up to 14 as some tasks are coded to more than one dimension

Table 5.16: Distribution of problem-solving tasks by technology dimensions

Dimension	No. *
Web	7
Spreadsheet	7
E-mail	4

*Does not add up to 14 as some tasks are coded to more than one dimension

Table 5.17: Distribution of problem-solving tasks by context

Dimension	No.
Personal	8
Work/Occupation	4
Civic	2

Table 5.18: Distribution of problem-solving tasks by intrinsic complexity (number of steps)

Dimension	No.
Single step	8
Multiple steps	6

Table 5.19: Distribution of problem-solving tasks by intrinsic complexity (number of constraints)

Dimension	No.
Single constraint	7
Multiple constraints	7

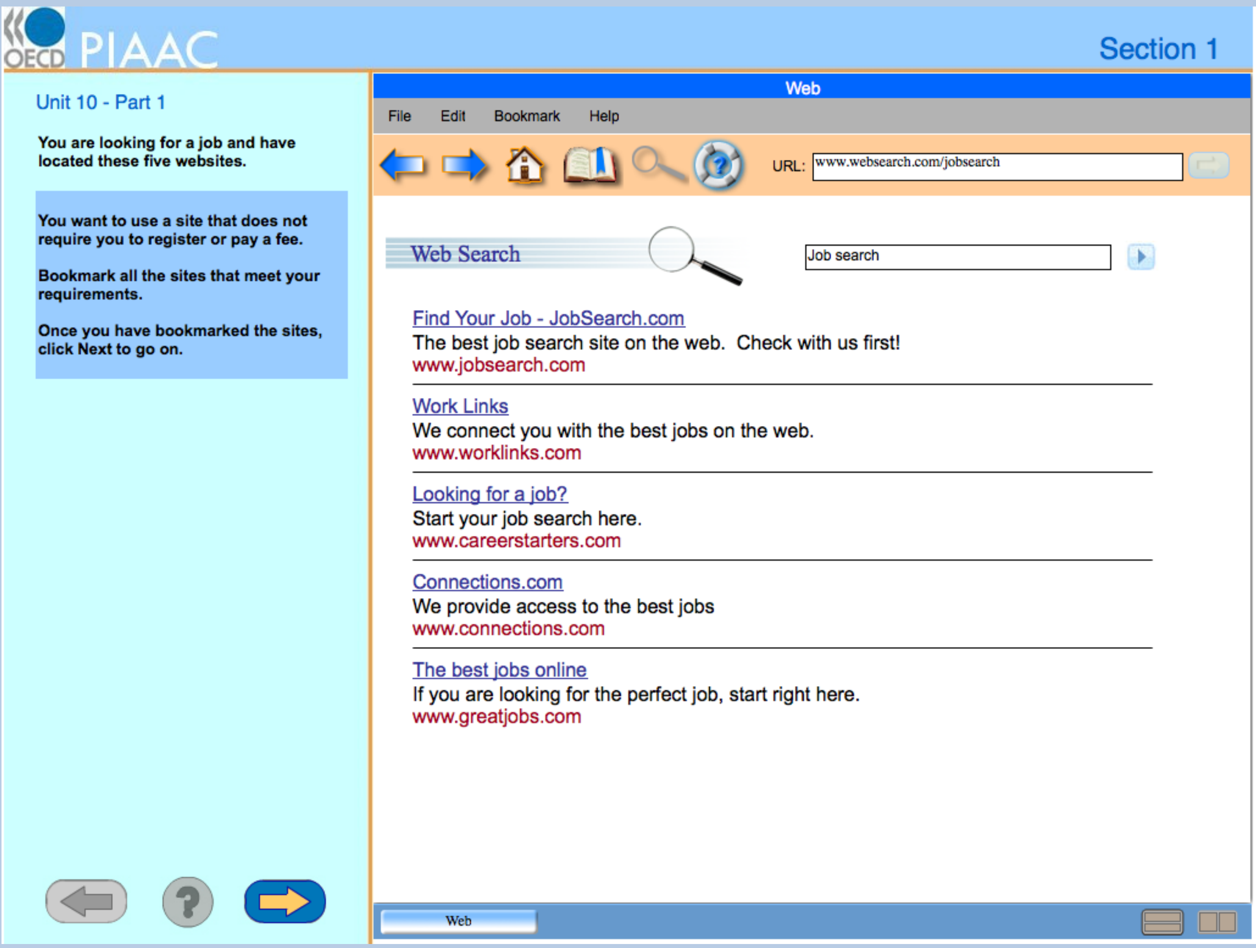
Table 5.20: Distribution of problem-solving tasks by explicitness of problem statement

Dimension	No.
Ill-defined problem statement	7
Well-defined problem statement	7

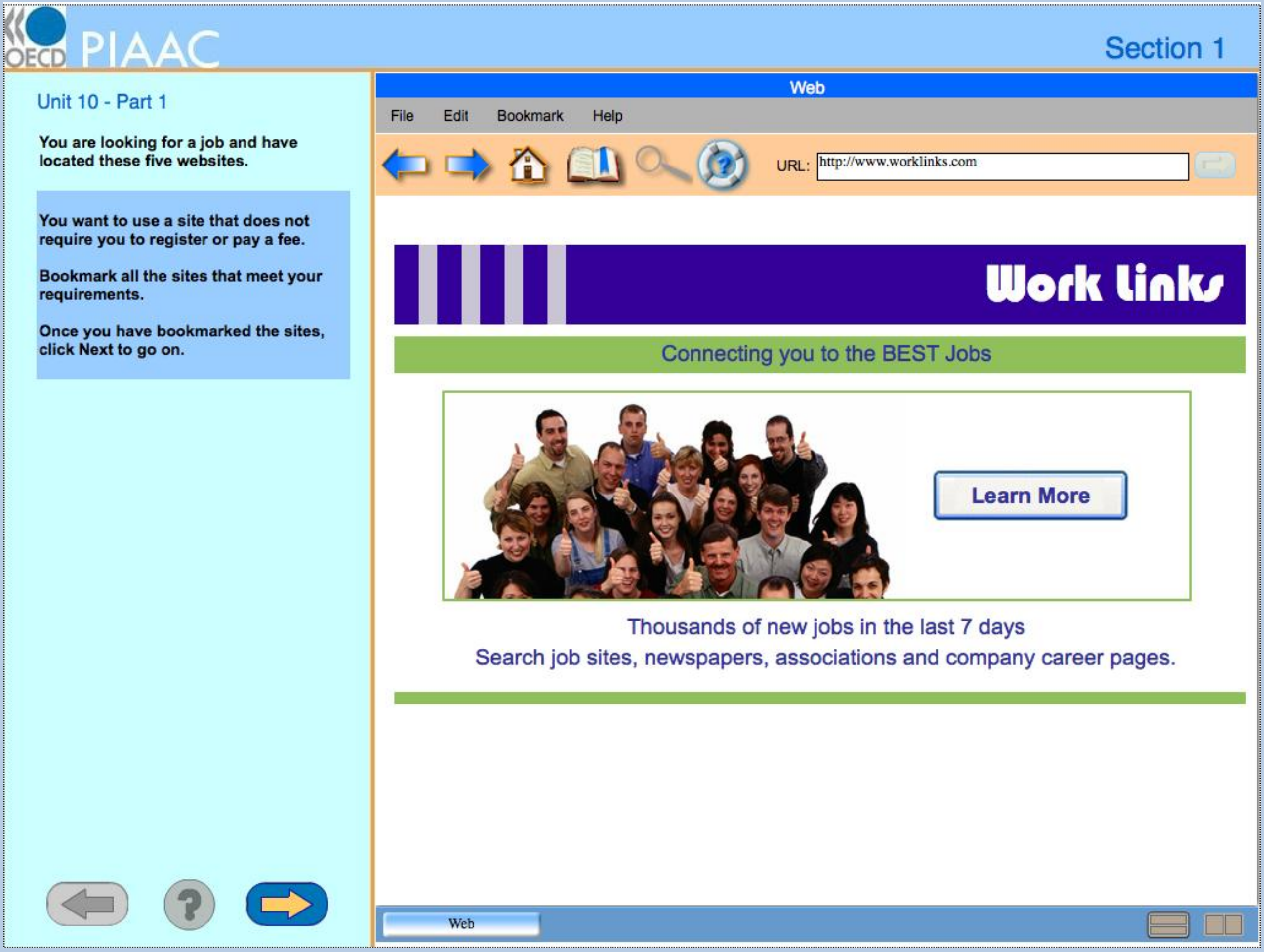
Problem Solving Sample Item

An example of a problem-solving item is provided below. This item involves a scenario in which the respondent assumes the role of a job-seeker. Respondents access and evaluate information relating to job search in a simulated web environment. This environment includes tools and functionalities similar to those found in real-life applications. Users are able to:

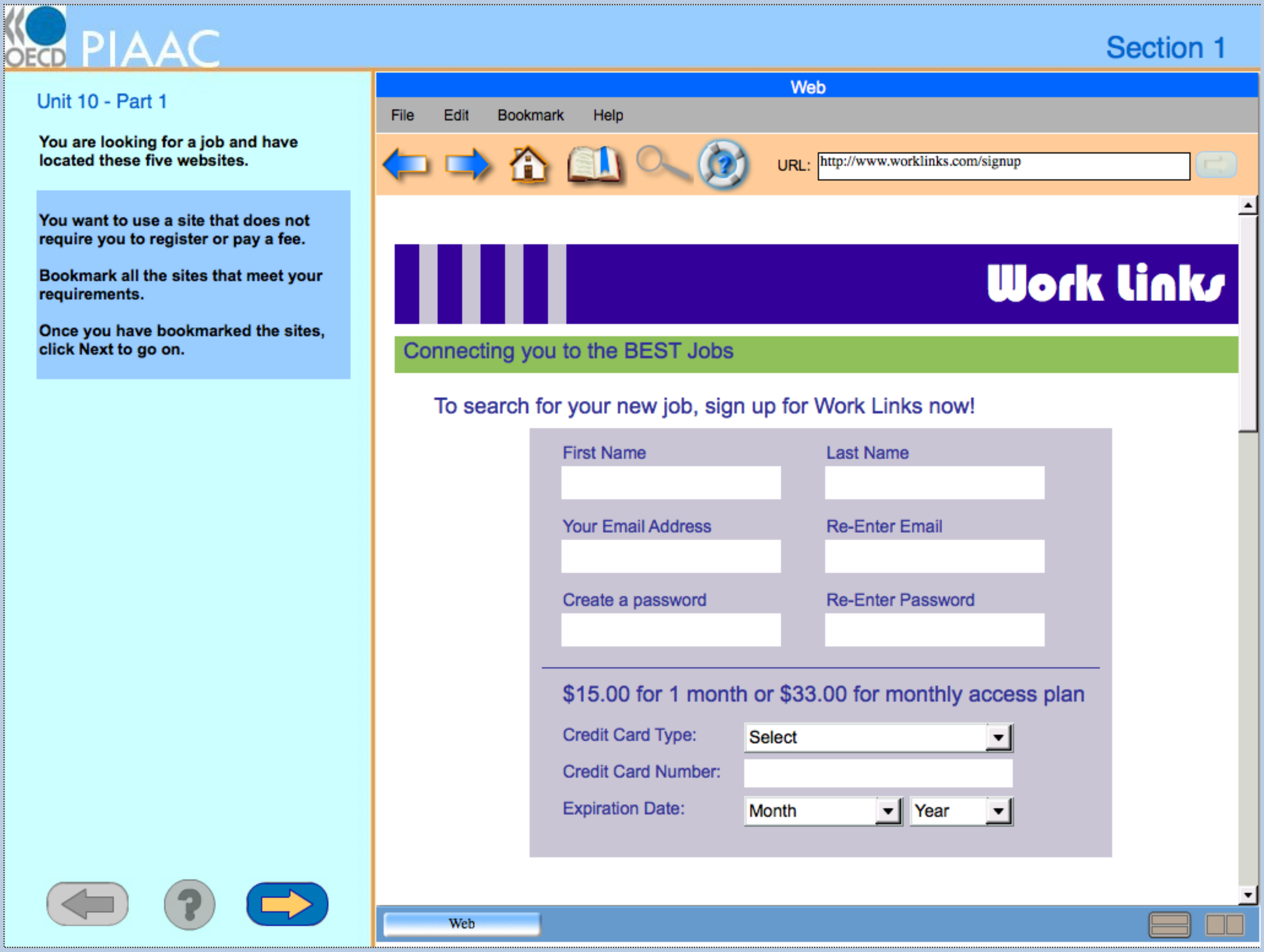
- Click on links on both the results page and associated web pages;
- Navigate, using the back and forward arrows or the Home icon; and
- Bookmark web pages and view or change those bookmarks.



The first stimulus accessed by respondents is the results page of the search-engine application, which lists five employment agency websites. To complete the task successfully, respondents have to search through the pages of the listed websites to identify whether registration or the payment of a fee is required in order to gain further information about available jobs. Respondents can click on the links on the search page to be directed to the websites identified. For example, by clicking on the "Work Links" link, the respondent is directed to the home page of "Work Links".



In order to discover whether access to the information on available jobs requires registration with the organization or payment of a fee, the respondent must click the “Learn More” button which opens the following page. The respondent must then return to the search results page to continue evaluating the sites in terms of the specified criteria, using the back arrows without bookmarking the page (correct answer) or having bookmarked the page (incorrect answer).



References

Organisation for Economic Co-operation and Development (OECD). 2013. *The Survey of Adult Skills: Reader’s Companion,* OECD Publishing. http://www.oecd.org/site/piaac/Skills%20(vol%202)-Reader%20companion--full%20v6%20eBook%20(Press%20quality)-27%2009%200213.pdf

6 Survey Methodology

Canada is a participant in the Programme for the International Assessment for Adult Competencies (PIAAC). The Canadian component was carried out in accordance with the standards in the PIAAC guidelines. These standards set out the minimum requirements for the survey design and the implementation of all phases of the survey, from planning to documentation.

6.1 Target population

The target population consists of all Canadian residents aged 16 to 65 inclusive, with the exception of long-term residents of collective dwellings (institutional and non-institutional), families of members of the Armed Forces living on military bases, and people living on Indian reserves. Because of operational constraints, sparsely populated regions were also excluded from the target population. Together, these exclusions made up no more than 2% of the total population of Canada, which easily met the international requirement that less than 5% of the target population be excluded from the survey.

Coverage of the survey's target population by the 2011 Census of Population was determined to be about 96% at the national level and between 94% and almost 100% at the provincial/territorial level (except for Nunavut). Table 6.1 shows preliminary estimates (as of March 2013) of the coverage rate of the population aged 15 to 64 based on 2011 Census coverage studies[6] for Canada and each province and territory. It should be noted, however, that the fact that someone was missed in the Census does not mean that he or she was also missed in the PIAAC, since Statistics Canada's interviewers had to prepare a roster of the members of the selected households before choosing the respondent.

[6]. Undercoverage estimates for the population aged 16 to 65 were not available at the time of writing.

Table 6.1 Estimated Census coverage rate of the population aged 15 to 64, Canada, provinces and territories, 2011

Region	Census coverage rate %
Canada	**96.4**
Newfoundland and Labrador	97.0
Prince Edward Island	96.1
Nova Scotia	96.4
New Brunswick	98.6
Quebec	97.8
Ontario	96.1
Manitoba	97.4
Saskatchewan	96.6
Alberta	94.9
British Columbia	96.0
Yukon	94.1
Northwest Territories	94.6
Nunavut	91.5

Source: 2011 Census Reverse Record Check, 2011 Census Overcoverage Study, preliminary results, March 2013.

6.2 Sampling frame

The response databases of the 2011 Census of Population and Housing and the National Household Survey (NHS) were used as sampling frames to construct the PIAAC sample.

These databases provided recent information about dwellings' usual residents so that people who are members of the survey's target populations could be selected. The Census was used for the general sample, the 16-to-24 age group in British Columbia, and linguistic minorities. NHS data were used to identify recent immigrants, Aboriginal people and Métis people. Only dwellings of Census or NHS respondents and dwellings whose residents were members of the target populations according to Census or NHS data were considered.

6.3 Sampling plan

A multi-stage probabilistic sampling plan was used to select a sample from each frame. The design produced sufficiently large samples for both official languages (English and French). In addition, the sample size was augmented to produce reliable estimates for a number of population subgroups, including young people (the 16-to-24 age group in British Columbia), linguistic minorities (Anglophones in Quebec and Francophones in New Brunswick, Ontario and Manitoba), immigrants who had been in Canada 10 years or less (i.e., since 2002), urban Métis in Ontario and urban Aboriginals.

In the territories, the initial sample was designed so that the final sample would contain at least 450 aboriginals in Yukon and Northwest territories and 600 in Nunavut. Note that initially, aboriginals in the territories were not explicitly targeted as such using their answers to the NHS, but households were stratified and sample sizes calculated in such a way that a sufficient number of aboriginal individuals would be interviewed to produce reliable estimates in each territory. As collection was conducted however, reports showed that these targets would not be met in Yukon and Northwest Territories. As a consequence, the initial sample in Yukon has been replaced by another random sample selected among NHS responding households explicitly targeting aboriginals according to the same criteria used in the provinces. In Northwest Territories, a portion of the sample selected in Yellowknife has been replaced by a random sample selected in communities known to have a higher percentage of aboriginals in their population.

In the provinces,[7] the primary sampling units (PSUs) were defined by updating the PSUs constructed for the 2003 International Adult Literacy and Skills Survey (IALSS).

At the time, Statistics Canada's Generalized Area Delimitation System was used to create PSUs with a sufficiently large population based on the number of dwellings within limited, reasonably compact areas. A general indication of the population's level of education according to the 1996 Census had been added to generate PSUs that reflected the distribution of levels of education in the province.

Since the enumeration area geography used in the 2001 Census was replaced, additional work was required to define the boundaries of each PSU in terms of dissemination areas before stratification and selection.

Using these boundaries and exclusions similar to the IALSS exclusions, the PSUs were allocated to the following strata: A (urban), B (rural) or E (excluded). PSUs were excluded when they were too large, did not have enough residents or were too far north. Reserves were also excluded. Further clean-up resolved cases of PSUs that were in more than one stratum. A few PSUs were divided or combined with others so that they would have an area and number of dwellings comparable to other PSUs.

In addition, 2006 Census data and the PIAAC's sample and target population sizes were used to update the stratum boundaries. Communities were formed to derive these stratum boundaries using dissemination areas or urban areas, depending on whether it was in a census metropolitan area (CMA) or the area of the

[7] In the territories, a two-stage sample design has been used. As a consequence, PSUs are constituted by households, and not by geographical areas.

CMA or urban area was greater than 5,625 km^2. The 2006 Census long questionnaire (2B) counts and the PIAAC's final sample sizes were also used to divide the communities into an urban stratum (A) and a rural stratum (B). The sample was divided among the PSUs on a preliminary basis using a Neyman allocation. Communities in which at least 15 dwellings had been selected were assigned to the urban stratum.

Stratification was then completed by assigning some PSUs to a new stratum, C, for which they were selected with certainty because of their size. The PSUs chosen for this stratum were those in which at least 80 dwellings had been selected for the general and special samples taken together, or in which 40 dwellings had been selected for a subsample.

After the final stratification was determined, a sample of PSUs was selected at the first stage in the rural stratum by sampling with probability proportional to the number of eligible persons in the PSU. In each province, the sample was distributed among the strata in proportion to actual population size, with a conservative design effect of 2.0 for the rural stratum and 1.5 for the urban stratum. The latter adjustment was made to compensate for the effect that the multistage sample design has on the variance of the estimates produced with the survey data.

In the urban stratum, the number of dwellings was estimated by allocating the initial sample size to strata A and C on the basis of the proportion of the general sample or the subsample for that PSU. In the rural stratum, the same sample size was allocated to all PSUs in the sample to equalize collection workloads.

In the urban stratum in provinces, as well as in the three territories, two-stage sampling was used. In the first stage, households were selected systematically with probability proportional to size. Size was defined as the number of adults aged 16 to 65 in a household based on 2011 Census data, at any time during the PIAAC collection period. The upper limit was set at four eligible adults for the core sample and three for the supplementary samples. In the second stage, the computer-assisted personal interview (CAPI) application used a simple random sampling algorithm to select one person from the roster of eligible adults that the interviewers made for each household during collection.

In the rural stratum, three-stage sampling was used. In the first stage, PSUs were selected with probability proportional to the number of adults aged 16 to 65 according to the 2011 Census. In the second and third stages, the selection method was the same as the one used for the urban stratum.

6.4 Sample size

The PIAAC sample was constructed from a general sample of 5,400 units, which were distributed among the provinces using a Kish allocation (Kish 1976) to obtain a sample of at least 5,000 English-speaking respondents at the national level. Then, 3,600 units from Quebec were added to produce a sample of 4,500 French-speaking respondents (required to meet the international consortium's standards). Supplementary units were added to this sample to produce more precise estimates for some provinces and territories and some subpopulations of interest.[8] Following adjustments for expected non-response and target population mobility, an overall sample of nearly 50,000 units was obtained. The samples were selected one by one in sequence, following the core sample. After each sample was selected, the

[8]. As noted previously, these supplementary units were added to meet the needs of federal, provincial and territorial government departments and ministries.

households chosen from the frame were removed before the next selection process, which made the samples dependent. Sequential selection of several samples in the same province or territory can be considered multiphase sampling.

In the final stage before sample selection, the size of the primary samples was augmented to compensate for a 6% vacancy rate among the selected dwellings and a 4% rate of households with no eligible members for the general sample, which made for a combined rate of approximately 11%.

The supplementary samples covered populations with specific characteristics, and because of natural mobility, a household selected for inclusion in one of these samples was more likely to have no eligible members at the time of contact with the interviewer, compared with the general sample. For example, persons aged 16 to 65 who moved out of a dwelling selected in the general sample shortly after the Census are very likely to have been replaced by other persons in the same age group; however, recent immigrants in that age group are less likely to have been replaced by other recent immigrants before the PIAAC was conducted. For this reason, the percentages used for the supplementary samples were different from the percentage used for the general sample. For example, the combined rate of vacant dwellings containing no members of the target group for the official-language-minorities sample was set at 15% in New Brunswick and 20% in Quebec, Ontario and Manitoba. A 65% unique response rate was also assumed, along with an 8% unique rate of refusal to share.[9]

Table 6.2 shows the expected number of 2012 PIAAC respondents by sample type for Canada and the provinces and territories.

---

[9]. The rate of refusal to share is the proportion of persons who responded to the survey but withheld consent for the transmission of their responses to organizations other than Statistics Canada and to the organizations responsible for processing the data collected. Those persons are treated as non-respondents.

Table 6.2 Expected distribution of PIAAC respondents by sample type, Canada, provinces and territories, 2012

Region	Expected Number of PIAAC Respondents		
	General sample	Supplementary samples	Total, PIAAC
Canada	**18,091**	**7,176**	**25,267**
Newfoundland and Labrador	1,399	0	1,399
Prince Edward Island	893	0	893
Nova Scotia	1,272	0	1,272
New Brunswick	1,098	368	1,466
Quebec	4,570	490	5,060
Ontario	2,635	2,530	5,165
Manitoba	922	1,225	2,147
Saskatchewan	913	600	1,513
Alberta	902	240	1,142
British Columbia	907	1,723	2,630
Yukon	900	0	900
Northwest Territories	900	0	900
Nunavut	780	0	780

Source: Programme for the International Assessment of Adult Competencies, 2012.

6.5 Supplementary samples

The supplementary samples were constructed with the Census or NHS response database. A dwelling could be included in one of these samples if data from the survey (the Census or the NHS) indicated that it contained at least one person with the desired characteristics. The criteria used for each supplementary sample are shown in List 6.1.

List 6.1 Definitions used to identify households containing persons of interest for the PIAAC supplementary samples, according to the Census or the NHS

Supplementary samples selected with census data	
Subpopulation of interest	**Definition**
Persons aged 16 to 24 (British Columbia)	Age according to Question 3 of census questionnaire 2A
Official language minorities (New Brunswick, Quebec, Ontario, Manitoba)	English: English is the person's only mother tongue according to Census Question 9 French: French is the person's only mother tongue according to Census Question 9
Supplementary samples selected with National Household Survey (NHS) data, stratum A (urban) only	
Subpopulation of interest	**Definition**
Immigrants in Canada 10 years or less (Quebec, Ontario, Alberta, British Columbia)	Persons who did not mark "Canada, by birth" (NHS Question 10) who have ever had landed immigrant status (NHS Question 11) and first became landed immigrants between 2002 and 2011 inclusive (NHS Question 12)
Métis (Ontario)	Persons who reported they were Métis and did not mark the "No, not an Aboriginal person" response in NHS Question 18
Aboriginals (Quebec, Ontario, Manitoba, Saskatchewan, Alberta, British Columbia), Replacement sample in Yukon	Persons who reported they were North American Indians, Métis or Inuit and did not mark the "No, not an Aboriginal person" response in NHS Question 18

At the time of the visit, the selected household was interviewed, and the interviewer checked that it was still eligible—i.e., that it had at least one person from the target population—using the same questions as the Census or the NHS. If more than one person was eligible, one of them was chosen at random. If the household was ineligible, it was coded as out of scope.

As a result, some households (for example, some of those selected in the Métis sample) were reported as out of scope because they no longer had any members with the desired profile.

Note that some members of the specific populations (recent immigrants, Aboriginals and so on) are also present in the sample of the general population, since they are members of the Canadian population aged 16 to 65.

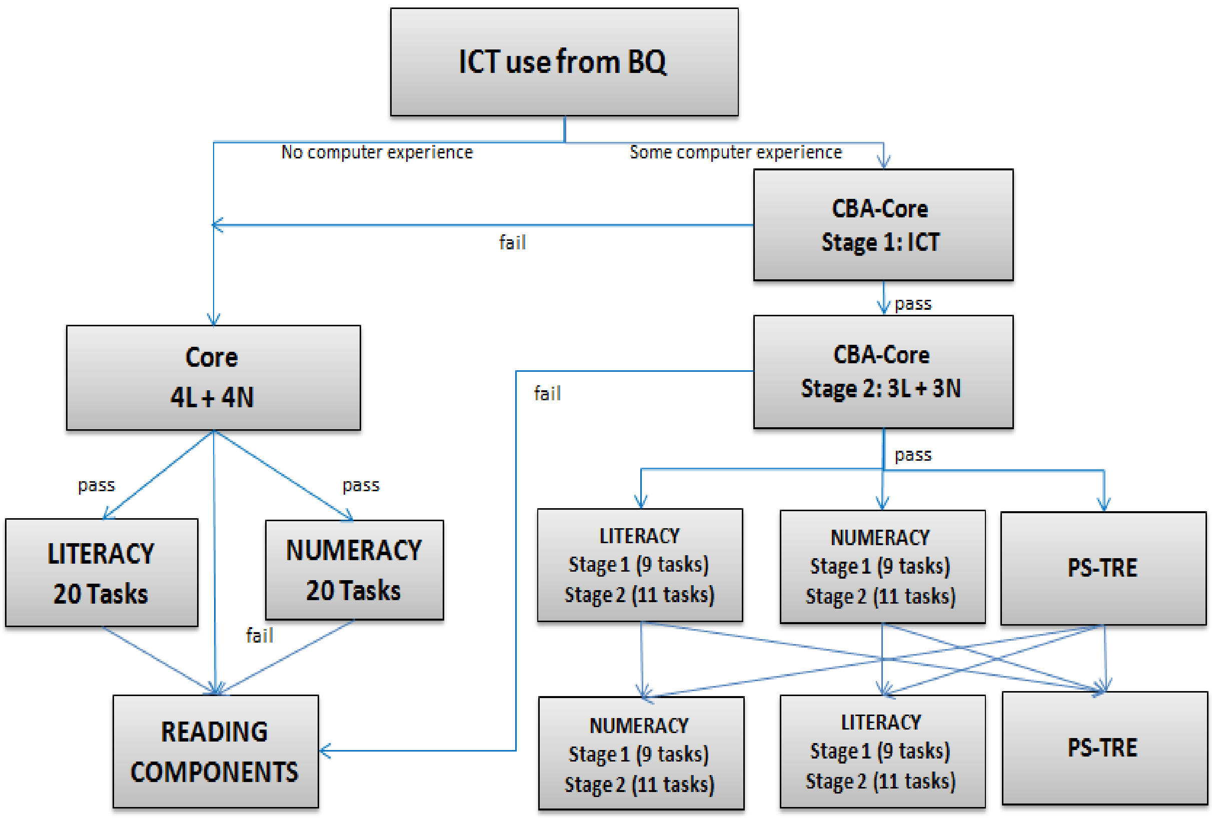
6.6 Data collection

6.6.1 PIAAC survey design, assessment design and application

The Programme for the International Assessment of Adult Competencies (PIAAC) is a survey of adult skills which is constructed of three main stages: the background questionnaire (BQ), the Core modules and the direct assessment part (direct assessment of literacy, numeracy and problem solving in technology-rich environments). While conceived primarily as a computer-based assessment (CBA), the option of taking the literacy and numeracy components through paper-based assessment (PBA) had to be provided for those adults who had insufficient experience with computers to attempt the assessment in CBA mode.

Respondents were initially asked to complete a set of basic questions about all household members, including their gender and age, in order to permit the random selection of one member from each dwelling. This “screener” collected as required, more demographic information aimed at identifying targeted sub-populations for the survey. The background questionnaire (BQ) was then asked of the selected respondent. The BQ included questions about respondents’ computer experiences, which were essential to branch them to either the paper or computer assessments at the end of the BQ. Respondents with no computer experience, based on BQ questions, and respondents who failed the Information and Communication Technology (ICT) core assessment were routed to the paper branch. Respondents with some computer experience also had the option to opt out of the CBA without attempting it and take the PBA. Most respondents, however, were routed to the computer branch of the survey. At the beginning of the survey, respondents were given the option of completing the survey in the official language of their choice (English or French). Prior to beginning the assessment, respondents were again asked in which of the official language they preferred to complete the assessment; from this point forward, respondents could not change their mind and had to complete the entire assessment in the language selected at that time. This necessitated a relatively complex design, which is presented graphically in the Figure 6.1 below.

Figure 6.1 Design of the PIAAC survey



Source: Programme for the International Assessment of Adult Competencies, 2012.

As seen in the above figure, there are several pathways through the assessment. Respondents with no experience in using computers, as indicated by their response to the relevant questions in the background questionnaire, were directed to the pencil and paper version of the assessment. Respondents with some experience of computer use were directed to the CBA where they took a short test of their ability to use the basic features of the test application (use of a mouse, typing, use of highlighting, and drag and drop functionality) – the CBA core Stage 1. Those who "failed" this component were directed to the pencil and paper pathway.

Respondents taking the computer path then took a short test (the CBA core Stage 2) composed of three literacy and three numeracy items of low difficulty to determine whether or not they should continue with the full assessment. Those who "failed" this module were directed to the reading components assessment. Respondents who passed this module continued on to take the full test and were randomly assigned to a first module of literacy, numeracy or problem solving items. Following completion of the first module, respondents who had completed a literacy module were randomly assigned to a numeracy or problem-solving module, respondents who had completed a numeracy module were randomly assigned to a literacy or problem-solving module, and respondents who had completed a problem-solving module were randomly assigned to a literacy, a numeracy or a second problem-solving module.

The assessment design assumed that the respondents taking the PBA path would be either those who had no prior experience with computers (as assessed on the basis of responses to the relevant questions in the background questionnaire) or those who failed the CBA core. It was, however, possible for respondents with some computer experience to take the PBA pathway if they refused the CBA.

Respondents taking the pencil and paper path first took a "core" test of four simple literacy and four simple numeracy items. Those who passed this test were randomly assigned to a module of either 20 literacy tasks or 20 numeracy tasks. Once the module was completed, respondents were given the reading-components test. Respondents who failed the initial "core" test proceeded directly to the reading-components test.

In Canada, the majority of respondents had enough computer skills to carry out the PIAAC assessment on the computer. Approximately 85% of respondents completed the Computer-based Assessment (CBA), and 15% completed the Paper-based Assessment (PBA).

The average times taken to complete the different stages of the PIAAC survey in Canada are as follows:

- Background Questionnaire (BQ): approximately 45 minutes;
- Paper-based Assessment (PBA): approximately 30 minutes;
- Reading Component Assessment: approximately 20 minutes;
- Computer-based Assessment (CBA): approximately 60 minutes.

6.6.2 PIAAC adaptive design

One of the unique aspects of the PIAAC was the adaptive design of the computer branch of the survey within the domains of literacy and numeracy.

Respondents were directed to different blocks of items on the basis of their estimated ability. Individuals who were estimated to have greater proficiency were more likely to be directed to groups of more difficult items than those who were estimated to be less proficient. Each of the literacy and numeracy modules was composed of two stages containing testlets (groups of items) of varying difficulty. Stage 1 contained three different testlets of nine items each, while Stage 2 contained four different testlets of 11 items each. Respondents' chances of being assigned to testlets of a certain difficulty depended on their level of educational attainment, whether their native language was the same as the test language (i.e. whether the language of the test was the first language or birth language of the respondent), their score on the literacy/numeracy core (CBA core Stage 2) and, if relevant, their score on a Stage 1 testlet.

Problem Solving in Technology Rich Environment (PS-TRE) is unique because of the nature of the domain; there was only one testlet per module. It was organized as two fixed sets of tasks: seven tasks in Module 1 and seven in Module 2.

Respondents directed to the paper booklet path directly started with a Paper Core booklet consisting of a set of items designed to determine whether they have the basic literacy and numeracy skills to proceed to the main assessment. This was scored by the interviewer, and if the respondent correctly answered a sufficient number of questions (4), they were then randomly assigned either a literacy or numeracy booklet.

Finally, PIAAC can provide more information about individuals with low proficiency levels by assessing reading component skills. This portion of the paper assessment was an international option and Canada was one of the participating countries. It measured basic reading skills using some short sections of exercises, word meaning, sentence processing, and basic passage comprehension.

With the exception of the reading components section (the time taken by respondents to complete the reading components tasks was recorded), no time limit was imposed on respondents completing the assessment, and they were urged to try each item whether it be on the computer or paper booklets. Respondents were given a maximum leeway to demonstrate their skill levels, even if their measured skills were minimal.

6.6.3 PIAAC quality control

To ensure high quality data, the international Technical Standards and Guidelines were followed and supplemented by adherence to Statistics Canada's own internal policies and procedures. The interviews were conducted in the respondent's home in a neutral, non-pressured manner. Interviewer training and supervision were provided, emphasizing the importance of precautions against non-response bias. Interviewers were specifically instructed to return several times to non-respondent households in order to obtain as many responses as possible. Extensive effort was expended to ensure that the home address information provided to interviewers was as complete as possible, in order to reduce potential household identification problems. Finally, the interviewers' work was supervised by using frequent quality checks at throughout collection and by having help available to interviewers during the data collection period. In total, Canada employed 786 interviewers over the duration of the survey.

The paper-based assessment was scored and captured in Statistics Canada. Explicit guidelines and a standard data capture tool were provided by the International Consortium to complete this work. As a condition of participation in the international study, it was required to capture and process files using procedures that ensured logical consistency and acceptable levels of data capture error. Specifically, complete verification of the captured scores (i.e., enter each record twice) was done in order to minimize error rates.

The International Consortium regarded Quality Control (QC) as an integral component to the overall success of the PIAAC survey. Various guidelines were established to ensure that the data collected by participating countries were reliable and valid.

The guidelines stipulated that throughout collection PIAAC countries routinely conduct validations to verify that an interview was indeed conducted or attempted as reported by the interviewer. Countries were required to validate at least 10 percent of each interviewer's finalized work to ensure that the case was handled according to study procedures. Validation included completed cases and those finalized with other outcome codes, such as vacant or refusal. Validation cases were selected randomly.

In Canada, the Quality Control Validation was done by a Computer Assisted Telephone Interview (CATI). The interview consisted of a series of questions about the respondents experience with the PIAAC survey, and the responses were then compared to the PIAAC survey data to determine if:

- the data matched (month and year of birth; education; address; demographics on household members; etc);
- procedures were followed (length of interview; composure of interviewer; interviewer using laptop; respondent completing assessment; interviewer helping respondent);
- the correct outcome code was assigned (correct vacant/ no contact/ absent/ seasonal dwelling etc).

If inconsistencies were discovered, the interviewer's entire completed caseload was then selected and subject to further validation in order to ascertain whether other cases were also compromised.

6.6.4 PIAAC coding

Industry, occupation, and education variables were coded using standard schemes such as the International Standard Industrial Classification (ISIC), the International Standard Classification of Occupations (ISCO) and the International Standard Classification for Education (ISCED). Coding schemes were provided for all open-ended items, as were specific instructions about coding of such items.

6.6.5 PIAAC data collection period

Data collection began in 2011 with the planning of interviewer assignments by the regional offices coordinating the collection activities. The first contacts with respondents were initiated in November 2011 across the country and the last interviews were completed in June 2012, with all survey-related materials being returned to head office by August of 2012.

6.7 Survey response and weighting

The Canadian PIAAC sample has a very complex design, involving stratification, multiple phases, multiple stages, systematic sampling, probability-proportional-to-size sampling, and several overlapping samples. It is also necessary to adjust for non-response at various levels. As a result, the estimation of population parameters and the corresponding standard errors depends on weighting coefficients, or weights. Two types of weighting coefficients were calculated: population weights, which are used to produce population estimates, and jackknife replicate weights, which are used to derive the corresponding standard errors.

6.7.1 Population weights

Since the PIAAC is a sample survey, each respondent was selected by means of a random process and represents a portion of the survey's target population. Each respondent's weight, i.e., the number of members of the target population that he or she represents, is calculated at the outset as the inverse of each person's probability of being selected in the sample. A sampling unit's overall probability of selection is the product of its probabilities of selection in all phases and stages of selection. The sequential selection of multiple samples in a province was taken into account by factoring in the probability that a unit selected in a given sample was not chosen in any previously selected samples. The initial weight was then adjusted to compensate for the various types of non-response in the survey.

There are four phases of weight adjustments for non-response: two apply to the weights before they are adjusted for the number of eligible members of the household, and two apply to the weights after that calculation.

For each type of weight adjustment, persons (respondents and non-respondents) with similar response probabilities were divided into response homogeneity groups (RHGs) for adjustment. For the adjustment of literacy-related non-response cases, the RHGs are composed of province–subsample combinations, because the number of literacy-related non-response cases in the sample is so small. For every other phase, in each province–subsample combination, an algorithm similar to the chi-square automatic interaction detection (CHAID) algorithm (Kass 1980) was used to form the RHGs. The RHGs were constructed so that each one had at least 30 households and a weighted response rate (or known eligibility rate for adjusting for the household's unknown eligibility at the household composition stage) of at least 40%.

The households selected in the sample were assigned to one of the following five response groups: respondent, literacy-related non-respondent (at this stage, only language problems were considered), non-literacy-related non-respondent, ineligible and unknown eligibility. They were allocated to the groups on the basis of the result codes selected by the interviewer when he or she contacted the people living in the selected dwellings and made a roster of the usual residents.

The first adjustment involves distributing part of the weight of dwellings of unknown eligibility among the dwellings that are ineligible (because they are vacant at the time of the interviewer's visit, they are being renovated, etc.). The second adjustment involves redistributing the weights of the dwellings of non-literacy-related non-respondents and ineligible dwellings among the weights of respondent dwellings.

After the roster of household members has been prepared and the respondent has been selected from the eligible members, a second code indicates whether the interview took place, and if not, why not. After the household composition stage, the members of a respondent household are in one of the following five response groups: respondent, literacy-related non-respondent, non-literacy-related non-respondent, ineligible member or disabled member.[10] The non-response adjustment stages that follow are applied to the weights, which reflect the number of eligible persons in the household.

The third non-response adjustment involves distributing the weights of disabled persons and selected non-respondents across the weights of respondents. Lastly, after the roster of household members is made, the fourth adjustment distributes the weights of pre-roster literacy-related non-respondents across the same type of non-respondents identified as persons selected to complete the survey.

Because of the overlap of the populations associated with the various samples, the weights had to be combined so that estimates could be produced using all units from all samples. The situation is similar to that of a survey with multiple frames, except that in this case, the samples are dependent. The weights were combined using the Hartley method (Hartley 1962) for multiple frames: The entire sample was

10. This category includes only persons whose disability, such as deafness or blindness, was considered incommensurate with participation in the survey.

allocated on the basis of the subpopulations targeted in the supplementary samples, and the weights were adjusted using coefficients proportional to the size of the various samples within the partition.

Lastly, the weights for each province and territory were calibrated separately using the calibration variables shown in Table 6.3.

The calibration totals used are population estimates based on the 2006 Census. They are official totals for the province, age and sex dimensions and simulation-based estimates for the other dimensions. Some missing data were imputed so that the variables used for calibration would be complete for all respondents.

Table 6.3 Calibration variables by province and territory

Province/Territory	Calibration Variables
Newfoundland and Labrador	Age group and sex, highest level of schooling, CMA in which the dwelling is located
Prince Edward Island	Age group and sex, highest level of schooling
Nova Scotia	Age group and sex, highest level of schooling, CMA in which the dwelling is located
New Brunswick	Age group and sex, highest level of schooling, membership in a linguistic minority
Quebec	Age group and sex, highest level of schooling, immigrant status, Aboriginal (all combined), CMA in which the dwelling is located, membership in a linguistic minority
Ontario	Age group and sex, highest level of schooling, immigrant status, Aboriginal, CMA in which the dwelling is located, membership in a linguistic minority
Manitoba	Age group and sex, highest level of schooling, Aboriginal (all combined), CMA in which the dwelling is located, membership in a linguistic minority
Saskatchewan	Age group and sex, highest level of schooling, Aboriginal (all combined), CMA in which the dwelling is located
Alberta	Age group and sex, highest level of schooling, immigrant status, Aboriginal (all combined), CMA in which the dwelling is located
British Columbia	Age group and sex, highest level of schooling, immigrant status, Aboriginal (all combined), CMA in which the dwelling is located
Yukon	Age group and sex, highest level of schooling, Aboriginal (all combined)
Northwest Territories	Age group and sex, highest level of schooling, Aboriginal (all combined)
Nunavut	Age group and sex, highest level of schooling, Aboriginal (Inuit)

Note: The age groups are 16 to 24, 25 to 34, 35 to 44, 45 to 54, and 55 to 65; the youth age group is defined as the 16-to-24 age group in British Columbia. Highest level of schooling can take four values: less than high school diploma, high school diploma, postsecondary education - below bachelor's degree, and postsecondary education - bachelor's degree or higher.

Source: Programme for the International Assessment for Adult Competencies, 2012.

The sample size and response rate for each province and territory are presented in Table 6.4.

Table 6.4 Actual sample size and response rate by province and territory, Canada, 2012

Region	Population aged 16 to 65	Initial sample	Out-of-scope cases[1]	Respondents[2]	Response rate[3] (%)
Canada	**23,381,067**	**49,450**	**6,335**	**27,285**	**58.5**
Newfoundland and Labrador	349,233	2,591	192	1,609	63.9
Prince Edward Island	97,542	1,656	135	929	59.5
Nova Scotia	627,538	2,361	129	1,441	60.4
New Brunswick	500,997	2,758	236	1,686	63.0
Quebec	5,404,254	9,699	842	5,911	62.5
Ontario	9,148,632	10,371	1,874	5,313	55.8
Manitoba	785,291	4,360	592	2,312	56.8
Saskatchewan	657,025	3,031	440	1,601	57.9
Alberta	2,622,199	2,211	226	1,224	56.8
British Columbia	3,111,300	5,376	1,066	2,733	58.8
Yukon	25,564	1,750	279	830	50.7
Northwest Territories	30,506	1,760	189	917	56.6
Nunavut	20,987	1,526	135	779	52.3

1. Out-of-scope cases are those that were coded as residents not eligible, unable to locate the dwelling, dwelling under construction, vacant or seasonal, or duplicate cases.

2. A respondent's data is considered complete for the purposes of the scaling of psychometric assessment data as long as the Background Questionnaire was completed.

3. Since the PIAAC sample has been selected among the 2011 Census or the NHS responding households, their respective response rates have been taken into to calculate the PIAAC weighted response rate.

Source: Programme for the International Assessment for Adult Competencies, 2012.

As required by the international consortium, two non-response bias analyses were carried out: a "basic" analysis, to assess the relationship between response status and available auxiliary variables correlated with the skill measures, and an "extended" analysis, to measure the effect of the various weight adjustments and assess the impact of non-response bias on key statistics (or correlated variables). These analyses showed that the various weight adjustments and the use of variables known to be correlated with the skill measures in the calibration stage minimized the effects that non-response had on the survey results.

6.7.2 Jackknife weights

A set of jackknife weights was generated to estimate the variance of the estimates produced with the survey data. The jackknife method with one unit removed (JK1) was selected because of its ease of implementation (Landry 2012). In the application of this method, each selected dwelling was assigned to a variance group. The sample PSUs were divided into 80 variance groups, or "replicates", and each replicate's jackknife weight was calculated by assigning a weight of 0 to the replicate's dwellings and multiplying the weights of the other dwellings by 80/79.

The method used to allocate the variance groups differs depending on whether the stratum is take-all (strata A and C) or take-some (stratum B). For a take-all stratum, the dwelling serves as the PSU, and each dwelling was assigned to a replicate independently. Thus, the first dwelling was assigned to a replicate at random, the next dwelling to the next replicate, and so on for all the dwellings in the stratum. The set of 80 replicates was split between the take-all PSUs and the take-some PSUs on the basis of a measure of the size (size of the PIAAC's target population) of the take-all or take-some PSUs. For example, if the take-all PSUs made up 50% of the PIAAC's target population, then 40 (80 * 0.5) replicates were allocated to the take-all PSUs. The remaining 40 replicates were assigned to the take-some PSUs. This process was performed independently for each province/territory–subsample combination.

Then the number of replicates to be allocated to each take-all PSU was determined so that the number of variance units assigned to each take-all PSU reflected the ratio of the PSU's size to a particular limit (the boundary between the take-all PSUs and the take-some PSUs). If a take-all PSU's size was about six times the limit, it received 6+1 replicates (i.e., six degrees of freedom). After the number of replicates was determined for each take-all stratum, the dwellings were sorted on the basis of the order in which they were sampled and the variance unit assigned to them. If the first take-all PSU in the sort received four replicates, its dwellings were assigned a variance unit of 1, 2, 3, 4, 1, 2, 3, 4, and so on. If the next PSU in the sort received two replicates, its dwellings were assigned a variance unit of 5, 6, 5, 6, and so on. The variance unit allocation for the take-all PSUs starts over when it reaches replicate n (in the example given above, replicate 40 would be followed by replicate 1).

The take-some PSUs were sorted into the order in which they were sampled. Then they were numbered sequentially from n+1 to 80 (in the above example, n would be 40) to form the variance units.

The presence of a second-phase sample among NHS respondents was also taken into account in the calculation of the jackknife weights by using the method described by Kim and Yu (2011).

The jackknife weights were produced from the PIAAC's entire initial sample, and the initial jackknife weights were calculated with the weights determined by the sampling plan. The entire weighting process was repeated for each of the 80 jackknife weights, including non-response weighting adjustments, combining of weights, and calibration.

References

Desjardins, Richard. 2004. *Learning for Well Being: Studies Using the International Adult Literacy Survey,* Institute of International Education, Stockholm University, Stockholm.

Green, D. A., and Riddell, W. C. 2007. *Literacy and the Labour Market: The Generation of Literacy and Its Impact on Earnings for Native Born Canadians*, Catalogue no. 89-552-MIE, no. 18, Statistics Canada. http://www5.statcan.gc.ca/bsolc/olc-cel/olc-cel?catno=89-552-MIE2007018&lang=eng

Hartley, H. O. 1962. "Multiple Frame Surveys", *1962 Proceedings of American Statistical Association, the Social Statistics Section*, pp. 203 to 206.

Kass, G.V. 1980. "An Exploratory Technique for Investigating Large Quantities of Categorical Data", *Applied Statistics,* no. 29, pp. 119 to 127.

Kim, J.K., Yu, C.L. 2011. "Estimation de la variance par répliques sous échantillonnage à deux phases", *Techniques d'enquête*, volume 37, pp. 73 to 81.

Kish, L. 1976. "Optima and Proxima in Linear Sample Designs", *Journal of the Royal Statistical Society*, Series A (General), volume 139, no. 1 (1976), pp. 80 to 95.

Landry, S. 2012. *Étapes de création des poids Jackknife en vue de l'estimation de la variance de l'EIA*, Statistics Canada, internal document.

Organisation for Economic Co-operation and Development (OECD). 2011. "PIAAC Main Study Virtual Machine", *Part I and II (Testing Procedures for the Cognitive Assessment)*, revised on 5 April 2011, pp. 6 to 10, internal document.

Organisation for Economic Co-operation and Development (OECD). 2011. *PIAAC Technical Standards and Guidelines,* December 2011, pp. 89 to 91, pp. 165 to 167. http://www.oecd.org/site/piaac/publications.htm

Organisation for Economic Co-operation and Development (OECD). 2012. *Literacy, Numeracy and Problem Solving in Technology-Rich Environments: Framework for the OECD Survey of Adult Skills*, OECD Publishing. http://dx.doi.org/10.1787/9789264128859-en.

Organisation for Economic Co-operation and Development (OECD). 2013. *OECD Skills Outlook 2013: First Results from the Survey of Adult Skills,* OECD Publishing. http://www.oecd.org/site/piaac/Skills%20volume%201%20(eng)--full%20v8--eBook%20(01%2010%202013).pdf

Organisation for Economic Co-operation and Development (OECD). 2013. *The Survey of Adult Skills: Reader's Companion,* OECD Publishing. http://www.oecd.org/site/piaac/Skills%20(vol%202)-Reader%20companion--full%20v6%20eBook%20(Press%20quality)-27%2009%200213.pdf

Organisation for Economic Co-operation and Development (OECD). 2013. *Technical Report of the OECD Survey of Adult Skills (PIAAC)*, OECD Publishing. http://www.oecd.org/site/piaac/All%20PIACC%20Technical%20Report%20final.pdf

OECD and Statistics Canada. 2005. Learning a Living: First Results of the Adult Literacy and Life Skills Survey. Catalogue No. 89-603-XWE. Statistics Canada. http://www.statcan.gc.ca/pub/89-603-x/2005001/4071714-eng.htm

OECD and Statistics Canada. 2011. *Literacy for Life: Further Results from the Adult Literacy and Life Skills*

Survey, Catalogue No. 89-604-XWE. Statistics Canada. http://www.statcan.gc.ca/pub/89-604-x/89-604-x2011001-eng.htm

Statistics Canada. 2005. *Building on Our Competencies: Canadian Results of the International Adult Literacy and Skills Survey*. Catalogue No. 89-617-XWE. Statistics Canada. http://www5.statcan.gc.ca/bsolc/olc-cel/olc-cel?catno=89-617-X&lang=eng.

7 Survey Procedures

7.1 Introduction

In many respects, the PIAAC procedures were guided by the international guidelines. Standard instruments, sampling, collection and processing methodology (including standardized code for Occupation, Industry and Education) are each important components in making PIAAC part of an internationally comparative study program. The following section outlines these procedures and details any deviations from the protocol in Canada.

PIAAC gathered descriptive and proficiency information from sampled respondents through a background questionnaire and a series of direct assessment (either on the computer or pencil-and-paper format) containing literacy, numeracy and problem solving (in technology-rich environments) tasks. Survey respondents spent approximately 45 minutes answering a common set of background questions concerning their demographic characteristics, educational experiences, labor market experiences, and literacy related activities. Responses to these background questions make it possible to summarize the survey results using an array of descriptive variables, and also increase the accuracy of the proficiency estimates for various subpopulations. Background information was collected by trained interviewers.

After answering the background questions, the remainder of respondents' time was spent completing the assessment portion (either on the computer or on a paper booklet) designed to measure their literacy, numeracy and problem solving skills.

To achieve good content coverage of each skill domain, the number of tasks in the assessment had to be quite large. Yet, the time burden for each respondent also needed to be kept within an acceptable range. To accommodate these two conflicting requirements—in other words, to reduce respondents' time burden without sacrificing good representation of the content domain—each respondent was administered only a fraction of the pool of tasks, using a variant of matrix sampling.

7.2 Model procedures manuals and instruments

Each PIAAC country was given a set of administration manuals and survey instruments to use as a model. Countries were permitted to adapt these models to their own national data collection systems, but they were required to retain a number of key features. For example, respondents were to complete the core and main assessment alone, without outside assistance. In addition, the use of the virtual machine (VM) software[11], as the main application to collect the survey results, was mandatory.

7.2.1 Background questions

The model background questionnaires given to each country contained two sets of questions: mandatory questions, which all countries were required to include; and optional questions, which were recommended but not required. Countries were not required to field literal translations of the mandatory questions, but were asked to respect the conceptual intent of each question in adapting it for use. Countries were

[11] The PIAAC VM consists in an Operating System (OS) fine tuned for PIAAC operations (lightweight and locked, including only the necessary third party software and system services) and a full platform running under this OS, which includes the background questionnaire and the assessment.

permitted to add questions to their background questionnaires (BQ) if the additional burden on respondents would not reduce response rates.

In order to facilitate the flow of the background questionnaire and to produce reliable and valid national results, Canada implemented structural adaptations to questions and response categories of the international BQ. These were adaptations that either changed the content of the question (e.g. asking two questions instead of one) or changed the content of the response categories (e.g. having 5 categories instead of 4). After collection, Canada had to derive the internationally required target variables from the nationally adapted or extended variables. These represent a total of 19 adapted national variables that were recoded to an international variable.

In Canada, PIAAC fielded all of the mandatory and optional questions, as well as a group of questions of national interest mostly related to mother tongue and immigrant or aboriginal status.

Below is a short description of the different BQ modules:

- **Section A** contains variables related to general information about the respondents, such as age, sex and language proficiency;
- **Section B** contains variables related to respondent's formal and informal education and training ;
- **Section C** contains variables on respondent's current status and work history;
- **Section D** contains variables related to the respondents current work experience and earnings;
- **Section E** contains variables related to the respondents recent/last work experience ;
- **Section F** contains variables related to use of different skills at work;
- **Section G** contains variables on use of literacy, numeracy, and ICT skills at work;
- **Section H** contains variables on literacy, numeracy, and ICT skills used in everyday life;
- **Section I** contains variables on attitudes and activities of the respondent; and
- **Section J** contains variables related to general household information, as well as questions about the respondent's background.

Where the answers to these questions do not compromise the confidentiality of our respondents, the PIAAC PUMF includes as much of the collected details as possible. Section 11 will examine the issues surrounding confidentiality in more details.

7.2.2 Assessment items

Like the International Adult Literacy Study (IALS) and Adult Literacy Life Skills Survey (ALL) before it, PIAAC is based on the premise that the difficulty of various literacy tasks is determined by certain factors, which are stable across language and culture. Accordingly, all of the PIAAC countries were given graphic files containing the pool of psychometric items and were instructed to modify each item by translating and adapting the master English text to their own language (English and French, in the case of Canada) without altering the graphic representation or task characteristics. Across countries, the instruments had to be equivalent in numbering and order of pages, layout of stimulus material and directives, graphics, response format, text format, and print quality or screen display. This consistency in the base materials minimized the effects of translation and adaptation errors.

As the intent of PIAAC is to have its results linked to previous international adult assessments, the assessment design assumes that 60 percent of the literacy and numeracy tasks will come from the ALL and the IALS. Therefore, these tasks must remain the same for PIAAC, particularly for countries that participated in these previous studies. Certain rules governed the item modification process. Countries were required to preserve conceptual associations during the translation process. Particular conventions used in the items—for example, currency units, date formats, and decimal delimiters—were adapted as appropriate for each country.

To ensure that the adaptation process did not compromise the psychometric integrity of the items, each country's items were carefully reviewed for errors of adaptation.

7.2.3 Standardized non-response coding

It was crucial that the PIAAC countries managed non-respondent cases in a uniform manner so as to limit the level of non-response bias in the resulting survey estimates.

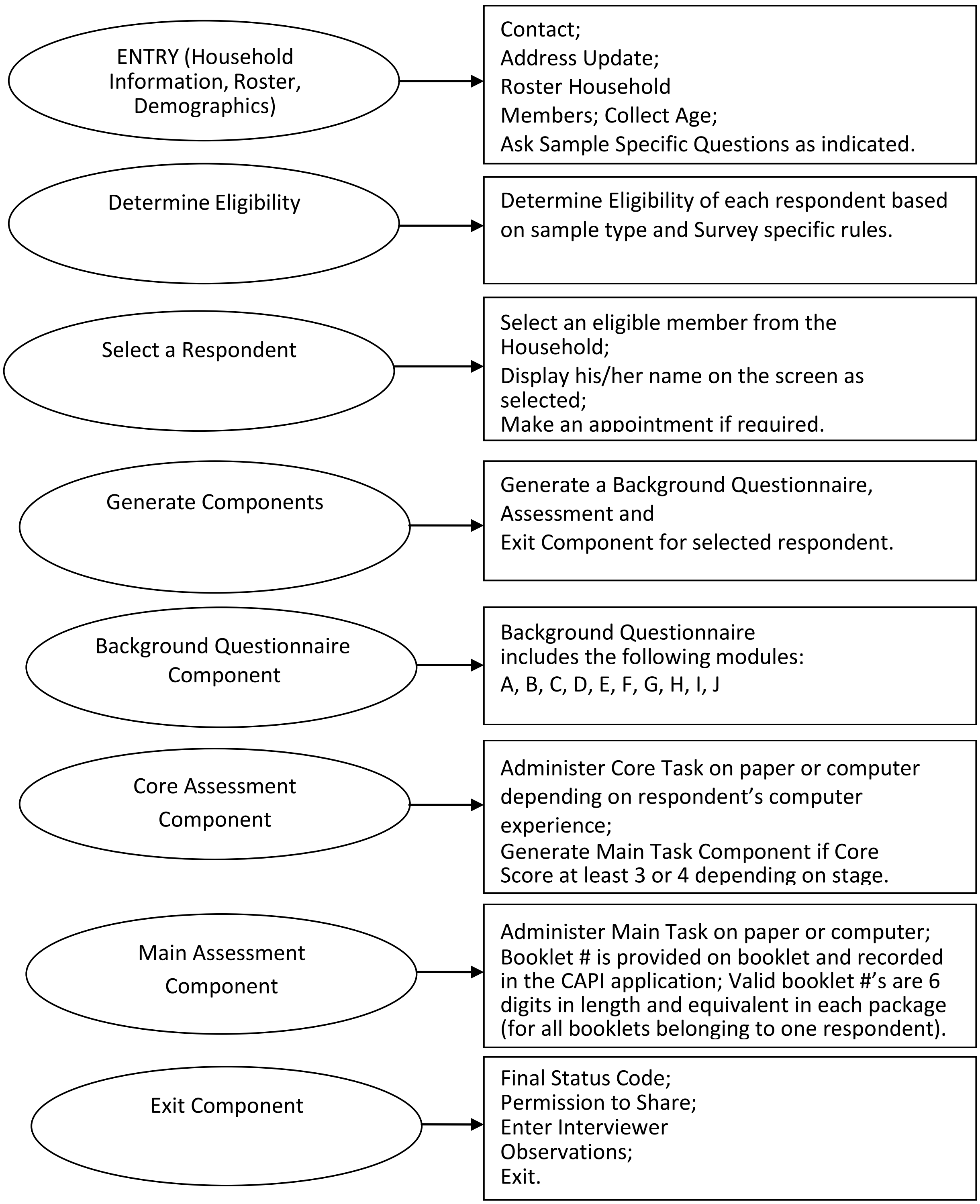
Proficiency values of PIAAC were calculated based on the cognitive responses as well as responses to the background questions. The respondents who completed both BQ and cognitive items provided the relationships between background variables and proficiency values. Through incorporating background information, measurement errors were reduced for the estimates of the proficiency distributions of any group of respondents.

Each country was responsible for hiring its own interviewing staff. Thus, the number of interviewers, their pay rates, and the length of the survey period varied among the countries according to their norms and budgets. Each country was provided with training documents and presentations to be adapted and used in training interviewers.

In Canada, PIAAC was collected by experienced and new Statistics Canada interviewers using Computer Assisted Personal Interviewing (CAPI) technology.

The diagram on the next page graphically depicts the main collection design of the Programme for the International Assessment of Adult Competencies.

Programme for the International Assessment of Adult Competencies
Main Collection Design



7.3 Scoring

The overall performance of items from the assessment was evaluated during the field test. The field test was used to evaluate scoring procedures, including scoring standards and scorer training for paper-based instruments and automated scoring procedures for the computer-based instruments. Items that did not appear to be performing as expected were examined and either revised or replaced for the PIAAC main study.

Respondents' literacy proficiencies were estimated based on their performance on the cognitive tasks administered in the assessment. Unlike multiple-choice questions, which are commonly used in large-scale surveys and which offer a fixed number of answer choices, open-ended items such as those used in PIAAC elicit a large variety of responses. Because raw data is seldom useful by itself, responses must be grouped in some way in order to summarize the performance results.

The models employed to estimate ability and difficulty were predicated on the assumption that the scoring rubrics developed for the assessment were applied in a consistent fashion within and between countries. Several steps were taken to ensure that this assumption was met.

For the large majority of respondents who took the assessment in its CBA format, scoring was done automatically. Manual scoring was necessary in the case of respondents taking the PBA version.

7.3.1 Scoring of tasks: Computer-based instruments automated scoring procedures

The purpose of this section is to explain in detail the scoring procedures within the computer branch of the assessment, focusing on the CBA Core, CBA Module 1 and CBA Module 2:

The Core: The word "core" is used in PIAAC to refer to two different sets of basic skills. Below are the scoring procedures for the CBA Core stages:

- **CBA core Stage 1 (Basic computer skills):** In the computer branch, the CBA core Stage 1 focused on basic computer skills including clicking, typing, scrolling, dragging, using pull-down menus and highlighting – skills respondents needed to complete the CBA main assessment. Thus, this module considered whether the respondents completed the task and was scored based on the completion of the action rather than the correct content. For example, one of the tasks asked the respondent to select "May" from a pull-down menu. The task was scored correctly if he/she used the pull-down menu to select any month. Out of the six tasks, respondents had to complete at least four tasks to move to the next stage. That is, respondents had to receive a score of 4, 5 or 6 AND they had to complete the highlighting task. Respondents who failed to demonstrate the necessary basic computer skills were routed to the paper branch. A successful completion of the CBA core Stage 1 led respondents to the CBA core Stage 2.
- **CBA core Stage 2 (Basic literacy and numeracy skills):** CBA core Stage 2 in the computer branch was designed to ensure that respondents had the basic literacy and numeracy skills necessary to proceed to the main assessment. CBA core Stage 2 contained six items with a passing score of at least 3; respondents with a score of 0, 1 and 2 were routed to the paper branch. For example to get a score of "4" a respondent had to answer 4 out of the 6 items correctly. The score received in the CBA core Stage 2 was used as a variable determining the choice of the first and second Testlet (i.e. Stage 1 and Stage 2 testlets) within Literacy and Numeracy.

The Modules: The CBA main assessment assessed the domains of literacy, numeracy and problem solving. Each respondent took two modules (Module 1 and Module 2), which each included two stages; Stage 1 contained three different testlets of nine items each, while Stage 2 contained four different testlets of 11 items each.

For the computer branch, the selection of a domain (literacy, numeracy or problem solving) for the first module (Module 1) is random. After completing Module 1 (either the two testlets for literacy or numeracy or the problem-solving module), the respondent proceeded to Module 2; the selection between Module 1 and Module 2 was also based on random probabilities. As noted in section 6.6.2, each of the literacy and numeracy modules was composed of two stages containing testlets (groups of items) of varying difficulty. All items were scored automatically.

Below are the scoring procedures for the CBA Module Stage 1:

- **CBA Module Stage 1:** Respondents needed to answer the items of each stage of a given module to get a certain score. For instance, in literacy and numeracy, the possible values of the stage 1 score of a module (and the result of the answers to the related items) were 0, 1, 2, 3, 4, 5, 6, 7, 8, or 9. E.g. to get a score "7" a respondent had to answer correctly seven out of the nine items. The Stage 1 score is used as a variable determining the test assignment for Stage 2 within Literacy and Numeracy.

Please refer to Figure 6.1 for the PIAAC Survey design.

7.3.2 Scoring of tasks: Paper-based instruments scoring procedures

Persons charged with scoring received intense training on scoring responses to the paper-based items using the PIAAC scoring manual. To aid in maintaining scoring accuracy and comparability between countries, the PIAAC survey used an electronic bulletin board, where countries could post their scoring questions and received scoring decisions from the domain experts. This information could be seen by all participating countries, and they could then adjust their scoring. To further ensure quality, monitoring of the scoring was done in two ways.

First, a certain proportion of booklets had to be re-scored. A minimum of 600 sets of Core Booklet/Exercise Booklet 1 or Core Booklet/Exercise Booklet 2 had to have been double scored within each country. The first score was considered as the main score; the second was considered as the reliability score. In Canada 1,000 sets of English and 1,000 sets of French Core Booklet/Exercise Booklet 1 or Core Booklet/Exercise Booklet 2 were double scored. This accounted for about 43 per cent of the total amount of booklets scored. The structure of the scoring design involved rescoring a large portion of booklets at the beginning and middle of the scoring process to identify and rectify as many scoring problems as possible. The goal in PIAAC scoring was to reach a within country inter-rater reliability of 0.95 (95% agreement) across all items, with at least 85% agreement for each item. In fact, most of the intra- country scoring reliabilities were above 95 per cent. Where errors occurred, booklets were reviewed and problem questions associated with a systematic scoring error by a particular scorer were rescored.

Second, the Consortium developed a cross-country reliability study where a set of anchor booklets were used to check the consistency of scorers across countries and to ensure they were applying the same criteria when scoring the items. The anchor booklets consisted of a set of 180 "completed" English booklets that were scored and rescored by every country.

Once Canada met the requirements of the two reliability studies (Canada had a within-country agreement above 97 per cent across items), the remaining Core, Exercise 1 and Exercise 2 booklets were single scored.

The section below explains the scoring procedures within the paper branch of the assessment, focusing on the paper core booklet (PPC), the literacy booklet (PP1), the numeracy booklet (PP2) and the reading components booklet (PRC):

- **PPC core (Basic literacy and numeracy skills):** the paper core booklet in the paper branch was designed to ensure that respondents had the basic literacy and numeracy skills necessary to proceed to the main paper-based assessment. The paper core contained eight items with a passing score of at least 4 (so scores 4, 5, 6, 7, and 8 were passing scores).
- In the **literacy booklet (PP1)** and the **numeracy booklet (PP2)** items were scored within Statistics Canada and each assigned a score of 1, 7 or 0. In general: A score of '1' was assigned for a correct response, a score of '7' was assigned for an incorrect response, and a score of '0' was assigned if no response was provided.
- The **Exercise Booklet RC (Reading Components**) was not scored within Canada; instead a procedure known as response capture was required. For each part of the Reading Components assessment, actual responses given by the respondent were captured in appropriate scoring sheets. During the data processing at the International Consortium, a response key was applied that assigned consistently coded scores for all reading component items. The following scheme was used: 0 = Question refused / not done, 1 = Correct response, 7 = Incorrect response and 8 = Any other response.

7.3.3 Paper-based within-country reliability study

During the main study within-country reliability study, participating countries were required to check the consistency of scoring by having a second scorer rescore a set of booklets taken by a sample of at least 600 respondents (in each national language) or 100% of the respondents (in each national language), whichever is less. As previously mentioned, 1000 sets of English and 1000 sets of French Core Booklet/Exercise Booklet 1 or Core Booklet/Exercise Booklet 2 were double scored in Canada. A scoring design was developed in Canada to accommodate our use of bilingual scorers on both English and French Booklets. In this design, once 1000 English booklet sets and 1000 French booklet sets were double scored, the remaining booklet sets were single scored and randomly distributed across all scorers.

The sample of respondents from which the set of booklets were rescored had to have been from respondents who took the Core paper booklet. That is, it had to include respondents who: i) entered the paper route because of lack of computer experience; or ii) failed the CBA Core Stage 1. Thus, these respondents must have been administered one of the following during the PIAAC survey: i) core booklet only; ii) core and literacy booklets; or iii) core and numeracy booklets. That is, respondents who failed the CBA Core Stage 2 and were administered only the reading components were not considered part of the within-country reliability sample set. This rescoring operation had to achieve a within-country inter-rater agreement of at least 95%.

The purpose of the within-country inter-rater reliability procedure was to ensure scoring reliability within a country and identify scoring inconsistencies or problems early in the scoring process so they could be resolved as soon as possible. In general, inconsistencies or problems are due to scorer misunderstanding of general scoring rules and/or a specific rule relating to a particular item.

The level of agreement between two scorers is represented by an index called inter-rater reliability. In PIAAC, inter-rater reliability represents the extent to which any two scorers agree on how a particular response should be scored, and thus the comparability in how the scoring rubric is being interpreted and applied.

Using data stored in the Data Management Expert (DME)[12], the IEA DME Tools Software was used for calculation of the inter-rater reliability. Once scores were entered into the database, the IEA DME Tools produced outputs and reports needed for examining scoring reliability.

Reliability and scoring errors were verified by reviewing the information for each item/scorer-pair in the Log/report generated by the DME tool. There were two rows/values in the Log/report that alerted the user to unacceptable reliability requiring further investigation. They were "Reliability flag (R)" and "N asymmetry (Nda)". The reliability flag alerted the user that the number of disagreements exceeded acceptable statistical limits and the item required further investigation. The "N asymmetry" value indicated if the disagreements were random (could usually be disregarded) or non-random (were investigated). Flagged items were reviewed and the scoring inconsistency was determined to be either natural or systematic.

If a systematic scoring error was suspected, the lead scorer investigated the scores of all booklets involved in the situation and looked for consistent scoring patterns which indicated the source of the disagreements. A few item pairs revealed systematic scoring errors. These were clarified with retraining, and all scores given by that scorer for that particular item were reviewed and corrected if necessary. Where no consistent score pattern was found for these disagreements, these disagreements were labeled as naturally occurring and no further action was required.

As well, during the beginning of the scoring process, the lead supervisor manually inspected a portion of the scored booklets for scoring accuracy before they were entered into the Data Management Expert software. This series of checks ensured that the initial booklets were scored according to the guidelines.

Scorers who received identical training within a country were expected to be more consistent with one other than with scorers in other countries. All of the rescoring reliabilities were above 97 percent. It is important to note that the results were well within the statistical tolerances set for the study. Table 7.1 displays the intra-country reliability for PIAAC 2012 in Canada.

7.3.4 Paper-based cross-country reliability study

Even after ensuring that all scorers were scoring consistently, clarifying ambiguities in the scoring guides, and correcting any systematic scoring errors, it was still necessary to examine the comparability of scores across countries. Accurate and consistent scoring within a country does not guarantee comparable scoring across countries. Scoring bias may be introduced if one country scores a certain response differently from the other countries. The cross-country reliability study described in this section was undertaken to ensure scoring comparability across countries.

[12] The Data Management Expert (DME) is a Microsoft SQL server based software developed by the Data Processing and Research Center (DPC) of the International Association for the Evaluation of Educational Achievement (IEA) in Hamburg, Germany. The use of the DME software and the delivery of the database in the DME's format were mandatory for all countries.

As noted earlier, responses to the PIAAC assessment items were scored by each country separately. To determine inter-country scoring reliabilities for each item, the Consortium developed a set of anchor booklets. The anchor booklets consisted of a set of 180 completed English booklets that were scored and rescored by every country, with no resolution of discrepancies. They included the international cover page and were identified by an international booklet serial number (or ID), prefilled in their cover pages. These anchor booklets were represented in Canada's scoring design after the first set of double scoring, as bundle 00. At least two bilingual scorers (fluent in French and English for Canada) had to score the English language international anchor booklets to ensure the equivalence of scoring across countries. The scores of these two bilingual scorers were compared and evaluated against the master scores for accuracy. The main and reliability scores for this bundle were used by the Consortium to calculate inter-rater agreement across countries.

Using the inter-country score reliabilities, researchers can identify poorly constructed items, ambiguous scoring criteria, erroneous translations of items or scoring criteria, erroneous printing of items or scoring criteria, scorer inaccuracies, and, most important, situations in which one country consistently scored differently from another.

Table 7.1 Scoring of paper-based Instruments: Within- and between-country agreement, Canada

Within-country agreement: Core (%)	Within-country agreement: Literacy (%)	Within-country agreement: Numeracy (%)	Within-country agreement: Reading Components (%)	Overall within-country agreement	Anchor Booklet Agreement (%)
99.4	97.3	98.4	99.8	98.7	97.3

Note: Bilingual scorers were used for all scoring within Canada.

Source: Programme for International Assessment of Adult Competencies, 2012

References

OECD. 2011. « PIAAC Main Study Virtual Machine ». *Part I and II (Testing Procedures for the Cognitive Assessment)*. Revised on 5 April 2011. Internal document.

8 Data Capture, coding and data processing

The Programme for the International Assessment of Adult Competencies (PIAAC) administers a psychometric assessment of adult skills in two modes (computer and on paper) in addition to the computer-assisted administration of a questionnaire. Given this design, the PIAAC data structures and formats are relatively complex and there is no single flat file as in previous surveys (i.e. in surveys such as ALL and IALS). For example, auxiliary and behavioral data will be collected and processed to support instrument validation, analysis, and reporting.

Because of these complexities and the timeline under which PIAAC was carried out, it was imperative to standardize, as much as practically possible, the procedures as they relate to the national and international data management.

For this reason, the Data Management Expert (DME)[13] software was provided to all countries and the delivery of the database in the DME's format was mandatory for all countries.

8.1 Data capture

The Canadian Background Questionnaire (BQ) and the majority of Assessments were completed on a computer. Approximately, 22% of the survey population completed the tasks on paper booklets, which needed to be scored and data captured, and, as part of the PIAAC guidelines, countries were advised to conduct complete verification of the captured scores (i.e., enter each record twice) in order to minimize error rates. Because the process of accurately capturing the test scores is essential to high data quality, 100 percent keystroke validation was performed.

8.2 Coding

Each country was also responsible for coding industry, occupation, language, country, region and education using standard international coding schemes (International Standard Industrial Classification, or ISIC; International Standard Occupational Classification, or ISOC; ISO 639-2; UN M49; OECD TL2; and International Standard Classification of Education, or ISCED). Further, coding schemes were provided for open-ended items, and countries were given specific instructions about the coding of such items so that coding error could be contained to acceptable levels. 20% of manually coded occupation and sector of industry data was verified by another coder. The average error rate for manually coded data could not exceed 10% for codes at the 4 digit level. As well, Canada checked the quality of the coding of the respondent's highest education level, occupation and industry against the distribution in the most recent Labour Force Survey or equivalent survey.

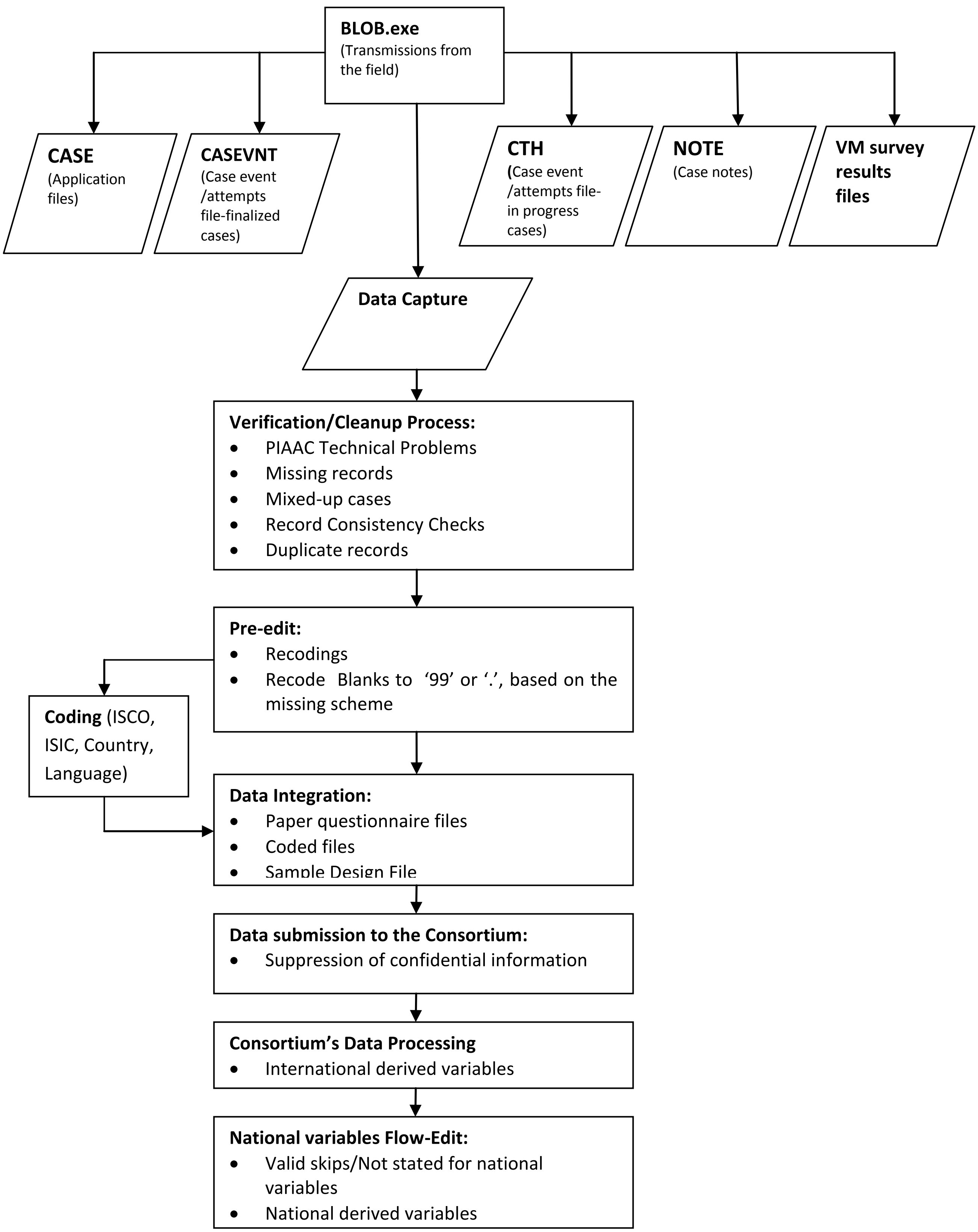
In addition, the Canadian file was also coded to National classifications for field of study (CIP), occupation (NOC) and industry (NAICS).

[13] The Data Management Expert (DME) is a Microsoft SQL server based software developed by the Data Processing and Research Center (DPC) of the International Association for the Evaluation of Educational Achievement (IEA) in Hamburg, Germany. The use of the DME software and the delivery of the database in the DME's format were mandatory for all countries.

8.3 Data processing

The following technical flowchart outlines the steps taken to process the Canadian data file as it arrived from the field.

PIAAC PROCESSING FLOW CHART



As illustrated above, the process can best be described as the culmination of the following 10 distinct steps:

1. Pre-collection Processing
2. Post-Collection Processing
3. The Verification/Cleanup Process
4. Pre-Edit
5. Data Integration
6. Imputations
7. Data submission to the International Consortium
8. Consortium's Data Processing
9. National variables Flow-Edit

Each of these will be described in greater detail in the sections below.

8.3.1 Pre-collection Processing (Step 1) – Data processing tools and database adaptation

The International Association for the Evaluation of Educational Achievement (IEA) in Hamburg, Germany is responsible for supporting the data management at the national level, as well as processing data at the international level. The IEA Data Processing and Research Center (DPC) developed a Microsoft SQL server based standard software (the "Data Management Expert" or DME) to integrate, manage and verify national databases. The use of the DME software and the delivery of the database in the DME's format were mandatory for all countries.

Along with the DME, countries were provided with the international codebooks. This was used as the template from which adaptations for national use are applied. The international codebooks contained all variables that were due to be submitted after data collection. These variables are assembled in datasets and the datasets contain certain instruments, if applicable, that were used for manual data entry.

The general goal was to include all international, as well as national data in Canada's database for submission. In order to do so, the international master codebooks needed to be adapted to reflect national instruments, i.e. all adaptations and extensions applied to Canada's BQ needed to be reflected in the DME as well. These adaptations may include the creation of national variables as well as the adaptation of valid ranges or value schemes of international variables.

Mark-all-that-apply questions were included in the mapping table as all responses given on such questions were presented under the same variable name in the PIAAC application. In the DME, these questions were split into as many variables as necessary. For instance, a 'Mark-all-that-apply' question with 8 different options resulted in 8 variables in the final DME database.

These national adaptations have been agreed upon submission, but the responsibility for implementing PIAAC at the national level, was a country responsibility.

8.3.2 Post-Collection Processing (Step 2)

8.3.2.1 Integration of the PIAAC application with the PIAAC VM

The complete PIAAC delivery system was made of two main components, the case management system (CMS) and the Virtual Machine (VM). The CMS was used to randomly select the eligible member from the selected household. It was also used to assist with the management of the cases during collection. The VM was the system built by the international consortium which included the background questionnaire, the overall workflow and the psychometric assessment. The background questionnaire and psychometric assessment part of the VM were adapted by each country to create a country specific system. In Canada, the system included both English and French components. An interview with a respondent was controlled by the workflow, which links to the various components.

Countries were required to use the VM for the PIAAC survey. Canada integrated the PIAAC survey systems (i.e. VM) with Statistics Canada standard case management system (i.e. Computer Assisted Interviewing Application – CAPI – programmed using Blaise). This was necessary in order for Canada to manage the Entry and Exit portions of the PIAAC survey, via their standard tools/applications (i.e. Blaise).

Below are some characteristics of the PIAAC application:

- Includes an Entry and Exit modules developed in Blaise.
- The PIAAC application Entry is used for the household rostering and used for contact with the selected respondent
- The household information related to the selected respondent completed in the roster of the Entry, will be transferred to the PIAAC application (i.e. VM) to continue on with the rest of the interview.
- The flow of the PIAAC application is the Entry, Content (VM) and the Exit. Included in the Exit is the Interview Observation Module (also known as the ZZ questions), which the interviewer had to complete once they leave the respondent's home.

8.3.2.2 The PIAAC application

The Collection steps gathered Screener information about the respondent and other members of the Household (i.e. age, gender, marital status, relationship to each other household member, etc). The Entry also had a section which asked questions about "First Language Learned at Home", Aboriginal status and Immigrant status. This section was to identify the special population sample and was only administered to selected households previously identified as meeting the criteria of one of these special populations. For every case, the interviewer had to complete the Entry section to permit the random selection of one member from each dwelling and to assign them to the PIAAC interview.

The content of the PIAAC application (VM) included the Background Questionnaire (BQ), Core and Exercises Modules (direct assessment of literacy, numeracy and problem solving in technology-rich environments). Respondents with no computer experience, based on BQ questions, and respondents who failed the Information and Communication Technology (ICT) core assessment were asked to complete a paper Exercise. Those who received the paper Exercise also had to complete a brief

Reading Component booklet and all respondents were asked a few exit questions, which were available in the Exit section of the PIAAC application.

The Exit section dealt with post-interview questions and procedures (i.e. assigning Outcome Codes) and marked the end of the interview. The PIAAC application Exit also included an Interview Observation Module, which was designed to collect information about the interview setting, the respondent’s behavior, and any events that might have interrupted or distracted the respondent during the Exercise. These are also referred to as “ZZ questions”. These questions were completed by the interviewer, as soon as possible after leaving the respondent’s home.

The entire interview was administered in the respondent's home by a Statistics Canada interviewer.

An overview of the PIAAC Application is presented graphically in the figures below (Figures 8.1 and 8.2).

Figure 8.1 Overview Flow Chart PIAAC Application

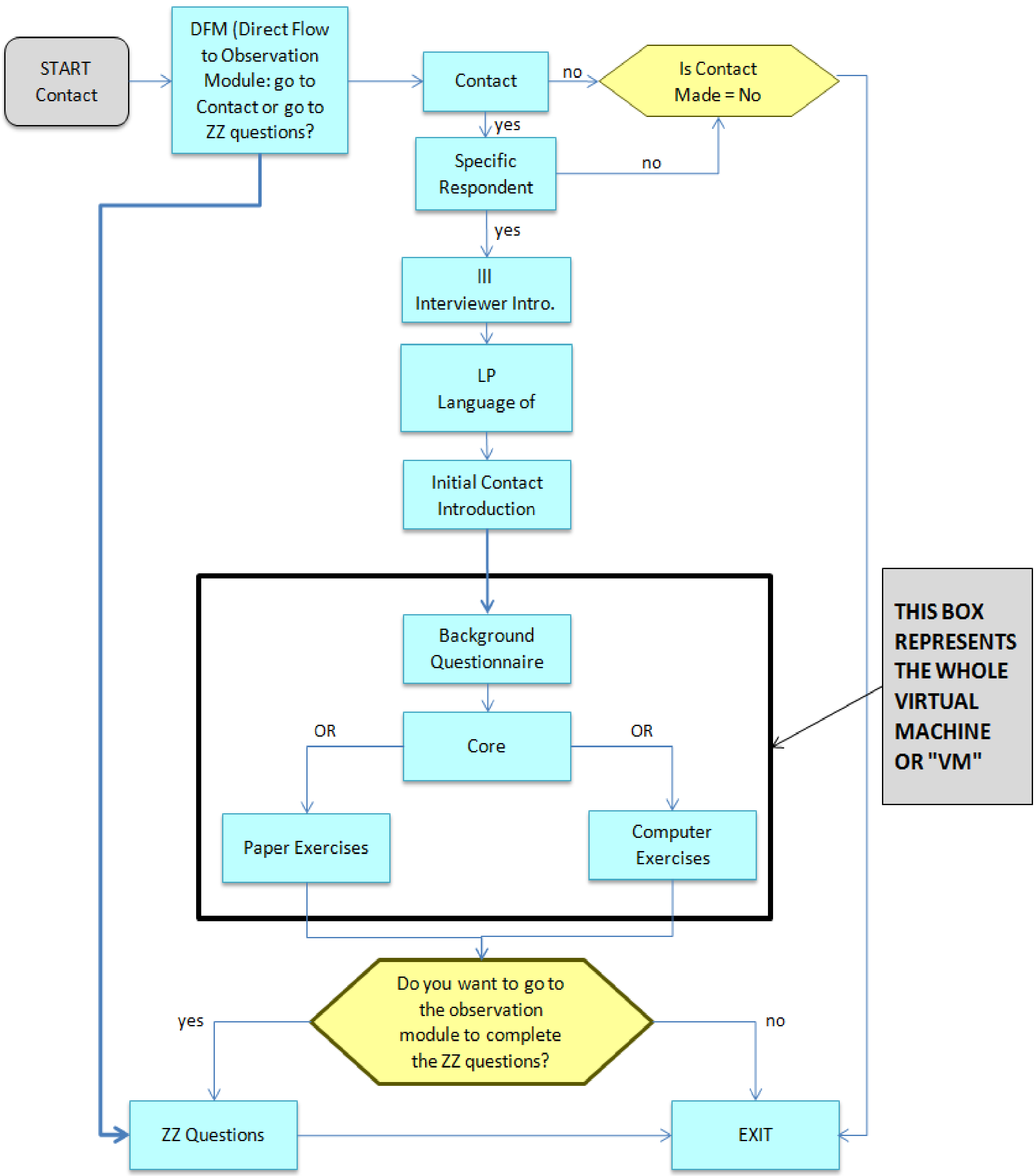
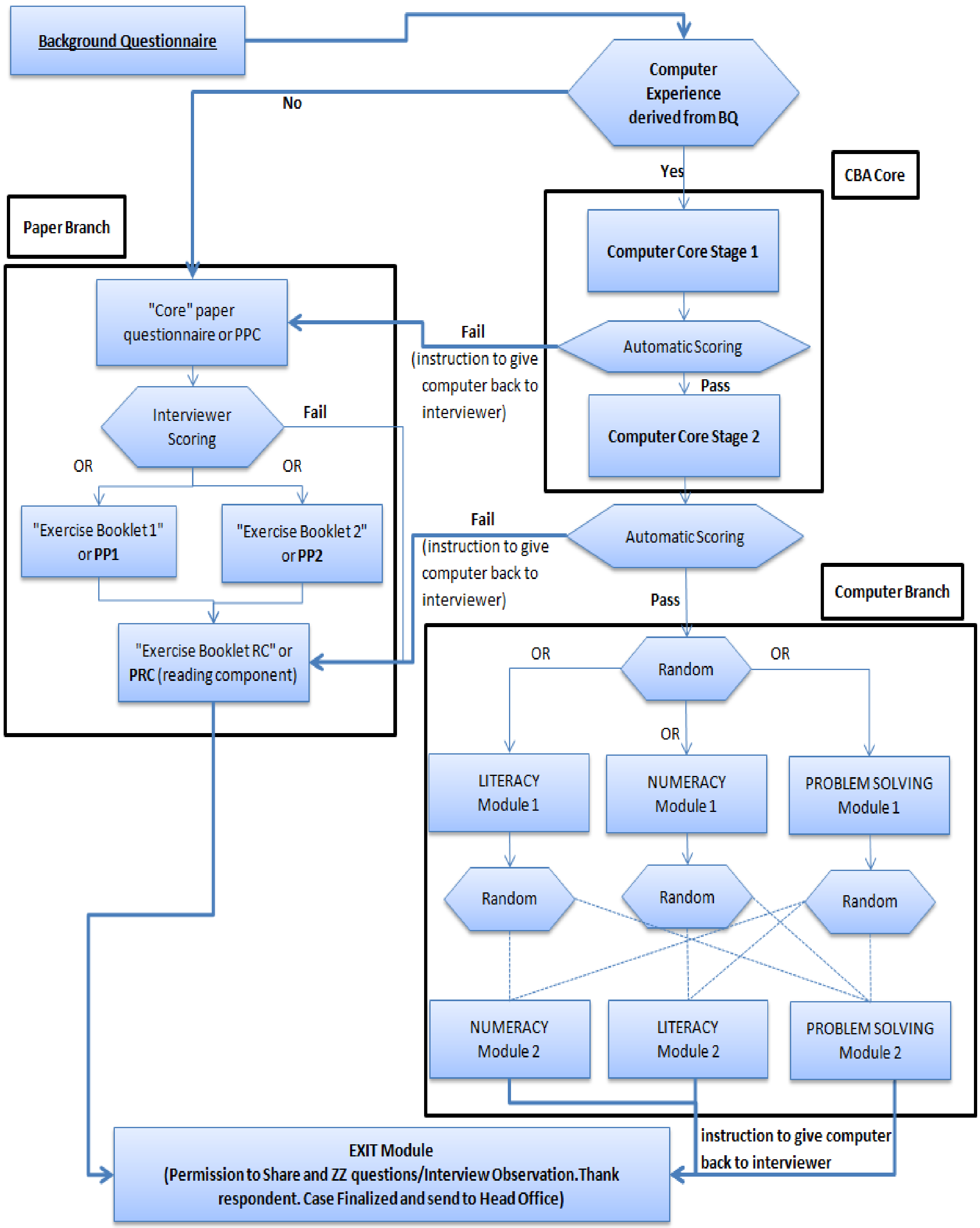


Figure 8.2 Overview Flow Chart of the Virtual Machine (VM) in PIAAC



8.3.2.3 Running the BLOB/Unpacking process

The BLOB process is a Statistics Canada process to dismantle data into various file structures as established by the record-layouts defined by the programmers who developed the Computer Assisted Interviewing Application (CAPI). The output of this process are a series of files such as the Header, Notes, Casevnt, CTH, PIAAC survey results (i.e. VM files) and other related files required for processing.

8.3.2.4 Importing the PIAAC survey results

The PIAAC survey results, created by the VM, had to be integrated into a national database using the DME software. This task consisted of importing the PIAAC survey results that were returned from the field and was performed on a daily basis.

8.3.3 The Verification/Cleanup Process (Step 3)

The purpose of the Clean up step is to identify any case that requires some type of intervention and identify cases that will not require editing because they are non-response. The process is intended to make corrections at the record level, such as:

- **identifying and resolving duplicate records** – for PIAAC only one record per sample unit can be included on a file. All duplicate records were thus either resolved or purged. There were two types of duplicate records: full duplicates (two identical records) and ID duplicates (two records with the same ID that contain different data);
- **verifying the BQ with the sample file** – the id captured in the Survey Data File was checked against the id in the sample file sent out for collection. Records received on daily basis were required to be a subset of the sample file. If such was not the case, an investigation was initiated and the appropriate corrective actions were taken. Reports were monitored on daily basis throughout the collection period;
- **identifying and resolving mix-up records** – interview completed at wrong household, not part of sample;
- **identifying and finding missing records** – missing from electronic transmission files or missing paper related packages/components;
- **identifying and resolving missing variables** – variables missing from specific survey files;
- **identifying and resolving technical problems** – cases that experienced technical problems/difficulties either in the PIAAC application (i.e. VM) or in the PIAAC Case Management System (CMS);
- **reviewing status code assignment** – status codes received from data collection were verified based on the information captured;
- **record Consistency Checks** – consistency edits examines the coherence of information between different variables. Many of the consistency edits were built into the PIAAC survey application (i.e. VM); values that do not concur with previous information was often flagged by the interviewer and clarified at the source. All of the required edits were repeated during the Canadian processing to ensure data quality. The PIAAC database (i.e. DME) also allowed countries to run pre-defined record consistency checks on the datasets contained in the database. The record consistency checks perform 45 checks that identify inconsistent records across all data sets.

- **excluding specific survey data – non-sharers and falsified cases**:
  a. Non-sharers: Canada's Case Management System (CMS) included a question that asked the respondents if they agreed to share, or not, their data at the International level. There are a total of 965 respondents that answered no, that they did not want to share their information at the International level. These are what we refer to as the non-sharers. These 965 respondents were excluded from the data.
  b. Falsified cases: These are cases that were deemed falsified after the quality control validation performed on a specific percentage of cases. These falsified cases were excluded from the data (see section 0 for more details on falsified cases).
- **identifying non-response cases.**

Once all records had been data captured and all steps identified in this section had been executed and all issues and problems resolved, the response, non-interviews and non-response files were prepared for processing.

8.3.3.1 Response Data File

The Survey Data File, in PIAAC, corresponds to the dataset BQR (Background questionnaire and global workflow – Results), found on the DME. This dataset comprises explicit, implicit or derived variables captured as part of the general workflow (random assignments, disposition codes, etc.); it includes all of the respondent's answers to the PIAAC questions, given while in the VM, either in the BQ or in the direct assessment modules.

There are a total of 28,599 cases that accessed the VM and that completed, either entirely or partially, the BQ and/or the direct assessment modules.

8.3.3.2 Non-interviews cases

Non-interviews are those where the roster was completed, a PIAAC respondent was selected, yet no interview took place due to refusal or other reasons. Consequently, such cases never made it into the VM. However, some information on the selected PIAAC respondent (age, gender, etc.) and outcome codes were captured by the PIAAC Entry section.

There are a total of 5,626 non-interview cases.

8.3.3.3 Non-Responses cases

Non-responses cases are those where data are missing. Households were sampled but eventually not screened and/or surveyed, for example because no contact could be established. In this case, only some sampling related information will exist.

There are a total of 15,262 non-responses cases.

8.3.4 The Pre-Edit (Step 4)

The Pre-Edit Process is performed after the Clean-up step. It consists of modifying the data at the individual variable level. Variables may be dropped, re-coded, re-sized or left as is.

8.3.4.1 Coding Other Specify

Questions that contain a list of answer categories often contain "other" as the final category. Coding these fields consisted of verifying the content and making a decision on whether the answer should have been included in another answer category or if they should stay in the "other" category.

In PIAAC, only variables that had an international requirement had their "other specify" category coded.

This process was done during the pre-edit stage; this meant that the 'write-in' responses were reviewed as part of a coding process.

8.3.4.2 Other pre-edit actions

Each variable passing through the pre-edit stage had one or more of the following actions taken as part of this processing step:

TERM	**DEFINITION**
Drop	Variable is removed from the file.
Re-code	Any code that may have been different in collection is coded to comply with International Standards (ex: recoding of national variables to international variables).
Leave	No modification to the variable is required.
Re-size	The length may need to be increased or decreased.
De-String	Mark-all that apply questions are returned in a compressed fashion, and need to be converted into "yes/no" values, one per possible answer.
Text	Written to Long answer file and removed from data file.

8.3.5 Data Integration (Step 5)

While there are some data that are being imported in the main processing database i.e. (DME) directly from the PIAAC survey results (i.e. VM files), other non-VM data needs to be integrated via generic export/import or manual data entry. This applies to:

- Scoring and response capture data from paper booklets from respondents;
- Scoring data from anchor paper booklets;
- Coding data such as ISCO and ISIC; and
- Sample design information.

All of these files were also subject to the same clean-up, verification and pre-edit steps outline above prior to being merged in the main processing database.

8.3.6 Imputations (Step 6)

Under international rules, imputation of missing values was not recommended. However in Canada some key variables had to be imputed to be used at the weighting step. These variables were age, age group, education level, year of immigration, Aboriginal status and mother tongue. A number of derived variables were created to support research into the antecedents and outcomes of skills. In most cases, these imputed values have not been included in the final data file. Generally speaking, imputation replaces missing values with plausible data; in some cases, however, the screener information and some other related information helped clarify and fix some of the inconsistencies in the values of these key variables. Such imputations were typically performed immediately following the consistency edits. Since some inconsistencies were dealt with in the preceding edit steps, imputation was kept to a minimum. When imputation was necessary, the donor imputation method was used: the value reported by a respondent with similar characteristics (the "donor") was copied to the record with the missing value. The donor was chosen at random from a group of potential donors. When no potential donors were available, the selection criteria were expanded. The successive criteria used to identify donors are listed in the paragraphs below (a), b), c), etc.).

Exercise language

The exercise language, which is used to assign the proper country code to the respondent (English Canada or French Canada), was generally derived from the exercise language (CILANG) and the language of the background questionnaire (BQLANG). For three cases, it was necessary to use donor imputation, based on province, mother tongue, language spoken most often at home and preferred official language.

Age of respondent

a) Primary source: If the respondent did not report his or her age, the age calculated from the respondent's date of birth in the background questionnaire (BQ) was used.
b) If the calculated age was unknown, the age reported in the screener was used, if it was consistent with the age group reported in the BQ. (There were 63 cases of this type.)
c) If the age was not available from sources a) and b), the age reported in the screener was used. (There were 28 cases of this type.)
d) If the age was not available from sources a), b) and c), the age was imputed with a donor record selected at random from respondents in the same age group and subsample. (There were 4 cases of this type.)

Gender of respondent

The gender question was not answered in 19 cases. The corresponding variable was imputed with information from the screener.

Education level

There were 93 records with no responses to the education level questions (variables B_Q01aca1, B_Q01aca2, B_Q01aca6, B_Q01aca7, B_Q01aca8, B_Q01eCA). The imputation model used values for other education questions, such as years of schooling or adult education experience. Imputation based on a suitable donor record was used to fill in this essential bit of missing information.

Born in Canada

BQ Question A_Q02CA (Were you born in Canada?) was imputed deterministically when the respondent was from the recent-immigrant subsample; it was assumed that the respondent was not born in Canada. (There were 4 cases of this type.) Otherwise, the response was imputed from a donor with the same values for province, age group, subsample and test language. (There were 26 cases of this type.)

Immigrant status and year of immigration

Question J_Q04bca5 (year immigrated to Canada) is required to distinguish between recent and established immigrants. Missing values for this variable and variable J_Q04bca3 (landed immigrant status) were imputed by selecting a donor at random from records with the same values as the recipient for the following variables:

a) province, subsample, exercise language, country of birth, age group, landed immigrant status (if this question was answered) and number of years lived in Canada (there were 35 cases of this type);
b) province, subsample, exercise language, country of birth, country of highest level of schooling, age group and landed immigrant status (if this question was answered) (there were 13 cases of this type);
c) province, subsample, exercise language, country of birth, age group and landed immigrant status (if this question was answered) (there were 12 cases of this type).

In all cases, the consistency of the imputed value was checked with the following edit rule: year of birth <= year of immigration.

Aboriginal status

All respondents were asked the Aboriginal status question. The question on membership in Aboriginal groups is a “mark all that apply” question. The questions on Aboriginal status and membership in the Métis group were imputed. For each record to be imputed, a donor was selected at random from records with the same values as the recipient for the following variables:

a) province, subsample, exercise language, country of birth, age group, landed immigrant status, Aboriginal mother tongue indicator and Aboriginal status (if this question was answered) (there were 50 cases of this type);
b) province, subsample, exercise language, country of birth, age group, landed immigrant status and Aboriginal status (if this question was answered) (there were 19 cases of this type);
c) province, subsample, exercise language, age group, landed immigrant status and Aboriginal status (if this question was answered) (there were 12 cases of this type).

Mother tongue

Mother tongue (Question A_Q03A1CA) is another variable that had to be imputed for the weighting step. For each record to be imputed, a donor was selected at random from records with the same values as the recipient for the following variables:

a) province, subsample, exercise language, country of birth, age group, landed immigrant status, preferred official language and language spoken most often at home (there were four cases of this type);
b) province, subsample, exercise language, country of birth, age group, landed immigrant status and preferred official language (there were 9 cases of this type);
c) province, subsample, exercise language, country of birth, age group and landed immigrant status (there were 20 cases of this type).

8.3.7 Submitting data to the International Consortium (Step 7)

After the integration and verification processes, each country was responsible for submitting the required materials to the PIAAC International Consortium.

Countries, such as Canada, may have regulations and laws in place that restrict the sharing of the main study data as it was collected with the Consortium and the OECD. A small number of directly identifying variables that are collected during the case initialization, were suppressed before submission to the Consortium. These are:

1. The respondent's first name;
2. The respondent's address; and
3. The respondent's telephone number.

8.3.8 Consortium's Data Processing (Step 8)

The Data Processing and Research Center (DPC) of the International Association for the Evaluation of Educational Achievement (IEA) in Hamburg, Germany is responsible for supporting the data management at the national level, as well as processing and cleaning data at the international level.

For the most part, the database files that were submitted to the IEA DPC have been transformed to match the desired target layout suitable for analysis, i.e. datasets were merged and the consistency of records across the initial datasets were checked. In addition, certain variables were modified, derived, or added.

8.3.9 National variables Flow Edit (Step 9)

8.3.9.1 National variables Flow Edit – Valid skips/Not stated for national variables

The Flow Edit replicates the flow pattern, (question order) intended by the survey design. The questionnaire includes a series of conditional branching depending on previously supplied answers. For instance, if the respondent indicates not ever taking any course or programs of study, the entire education section is skipped. Fields that are skipped because of flows are converted to "6, 96, 996, etc." (valid skip), as well fields that are skipped because of non- response are set to "9"'s (Not Stated).

In PIAAC's background questionnaire, a value is used to identify validly skipped questions (6, 96, 996, …), depending on the flow of questions in the international BQ. This value will not be assigned to national questions/variables as the processing of national adaptations and extensions was not the Consortium's responsibility. Missing values corresponding to "not administered" portions of the exercise items or missing values reflecting validly skipped questions in the background questionnaire were derived by the Consortium only for international variables. Assignment of valid skips and not stated to national variables was a country responsibility.

Missing value "Not stated / available / inferred" (9, 99, 999, …) is a systemic non-response and is assigned if a variable originally had a value but the value was out of range or otherwise useless and could not be reconciled or fixed. This value will be applied by the international data processing systems and/or by countries in some agreed cases.

Canada applied valid skips to the national variables using the processed Canadian data files provided by the Consortium, using a Top-down approach. The top-down approach follows the flow of the questionnaire exactly starting with the first question, therefore if the answer to the first question requires a skip of 3 questions then the 3 questions skipped would be set to valid skips (6, 96, 996, …) without regard to their contents.

Where necessary, missing value "Not stated / available / inferred" (9, 99, 999, …) may also be applied to the Canadian variables.

References

IEA Data Processing and Research Center (DPC). 2013. *PIAAC Main Study (MS) Data Processing Documentation*. Version 1.0. 31 Jan 2013. Internal document.

Organisation for Economic Co-operation and Development (OECD). 2013. *The Survey of Adult Skills: Reader's Companion,* OECD Publishing. http://www.oecd.org/site/piaac/Skills%20(vol%202)-Reader%20companion--full%20v6%20eBook%20(Press%20quality)-27%2009%200213.pdf

9 PIAAC Data Explorer (PDX)

The PIAAC Data Explorer (PDX) is a web-based statistical application intended for consumption by a broad audience of users interested in querying the PIAAC database. The PDX is built to be widely accessible via a web browser, and is based on industry-standard technologies. Because it is web-based, and processing takes place on a central server, the PDX can be accessed and used with computers that meet fairly simple requirements. The user's computer is used only to create a request or data query, deliver the request to a central server where processing takes place, and then receive and display back the results in a user friendly format.

A typical query consists of the user selecting the year, subject, jurisdiction(s), and variable(s) of interest. Then the user proceeds to select the statistics of interest and format the table. The user is able to select among the following statistics: average scale scores, percentage of respondents, standard deviations, percent reaching scale levels, and percentiles. These statistics are calculated for each of the subgroups defined by the variable or variables, selected either one variable at the time, or in cross-tabulation mode. In addition, the user is able to collapse categories for each of these variables and use the collapsed categories in the analysis. All statistics are calculated using weighted data, with their corresponding standard errors taking into account sampling and measurement uncertainty. The user has the option to select whether the standard errors are displayed in the table or not, as well as the precision with which the statistics are displayed. The results can then be displayed in a table or in a graphic.

Regardless of whether the results are displayed in a table or graphic mode, the results can be saved or exported for further post-processing or for inclusion in an external document. Export formats currently available include MS Word, MS Excel, PDF and HTML.

A significance test module allows the user to specify significance testing to be done between subgroup means, percentages and percentiles, within and across cycles, while implementing necessary adjustments that take into account the sample and test design, as well as adjustment for multiple comparisons. Significance test results can be displayed in table or in graphic format.

Table results can be easily exported to be manipulated using spreadsheet software, allowing the user to customize the titles and legends of the tables, and to do any required post processing. Likewise, the graphic results can also be exported to be included in documents for future display and use in reports and presentations.

Several methods are used to ensure confidentiality in the PIAAC Data Explorer:

- Personal information not on file
- Smallest geography unit is province
- No counts (unweighted or weighted population counts)
- Proportions rounded to unit (without decimals)
- Local suppression when unweighted count is smaller than 60.

References

ETS. 2013. *PIAAC Data Explorer (PDX) Functionality*. Internal document.

10 Guidelines for tabulation and analysis

This section summarizes the guidelines to be followed by users tabulating, analyzing, publishing or otherwise releasing any data from the survey microdata tapes. With these guidelines, microdata users should be able to produce figures similar to those produced by Statistics Canada as well as figures that have not yet been published in accordance with these guidelines.

10.1 Sample weighting guidelines for tabulation

Literacy surveys are based on complex sample designs, with stratification, multiple stages of selection, and unequal probabilities of selection of respondents. Using data from such complex surveys presents challenges for analysts because the survey design and the selection probabilities affect the estimation and variance calculation procedures that should be used. In order for survey estimates and analyses to be free from bias, survey weights must be used.

While many analysis techniques in statistical applications allow weights to be used, the meaning or definition of the weights differs from what is appropriate in a sample survey. As a result, while in many cases the estimates produced by the applications are correct, the variances they calculate are not very precise.

10.2 Definitions of types of estimates: categorical vs. quantitative

Before the methods of tabulating and analyzing the PIAAC data are considered, it is useful to describe the two main types of point estimates of population characteristics that can be generated from the survey's microdata file.

Categorical estimates

Categorical estimates are estimates of the number or percentage of members of the survey population that have certain characteristics or belong to a particular category. Examples of such estimates are the number of Albertans at Level 1 on the literacy scale and the proportion of Canadians at Level 4 on the numeracy scale. An estimate of the number of persons who have a certain characteristic may also be referred to as an estimate of an aggregate.

Examples of categorical questions:

Q: Do you ever watch television or videos in a language other than English or French?
A: **Yes / No**
Q: How would you rate your reading skills in English needed in daily life?
A: **Excellent / Good / Moderate / Poor**

Quantitative estimates

A quantitative estimate is an estimate of the total, the mean, the median or some other measure of central tendency of a quantity, based on some or all members of the survey population. It basically takes the form $\hat{X}/\hat{Y}$, where $\hat{X}$ is an estimate of the quantity's total for the survey population and $\hat{Y}$ is an estimate of the number of persons in the survey population who contribute to that total.

An example of a quantitative estimate is the average number of employers that Canadian workers had in the past 12 months. The numerator is an estimate of the total number of employers that Canadian workers had in the past 12 months, and the denominator is the number of Canadians who reported that they worked in the past 12 months.

Examples of quantitative questions:

Q: How many different employers have you had in the past 12 months?
A: **|_|_| employer(s)**
Q: How many hours per week did you usually work at this job?
A: **|_|_| hours**

10.2.1 Tabulation of categorical estimates

Using the microdata file, estimates of the number of people in a given country with a certain characteristic can be obtained by summing the final weights of all records possessing the characteristic of interest.

Proportions and ratios of the form $\hat{X}/\hat{Y}$ are obtained for a country by

a) summing the final weights of records having the characteristic of interest for the numerator ($\hat{X}$),
b) summing the final weights of records having the characteristic of interest for the denominator ($\hat{Y}$), and
c) dividing the numerator estimate by the denominator estimate.

10.2.2 Tabulation of quantitative estimates

Using the microdata file, quantitative estimates can be obtained by multiplying the value of the variable of interest by the final weight for each record, and then summing this quantity over all records of interest. For example, to obtain an estimate of the total number of different companies for which part-time workers in a particular country have worked in the past 12 months, multiply the value reported in Question C_Q10a (number of companies) by the final weight for the record, and then sum this value over all records with C_Q07=2 (part-time).

To obtain a weighted average of the form $\hat{X}/\hat{Y}$, the numerator ($\hat{X}$) is calculated by the quantitative estimation method and the denominator ($\hat{Y}$) by the categorical estimation method. For example, to estimate the average number of companies for which part-time workers in a particular country worked in the past 12 months,

a) estimate the total number of companies as described above,
b) estimate the number of people in this category by summing the final weights of all records with C_Q07 =2,
c) divide estimate a) by estimate b).

10.3 Skill level estimates

The PIAAC sample design is an adaptation of the three-parameter logistic model used in item response theory (IRT). The first parameter (A) is the item's discriminating capacity (sensitivity to proficiency) and the second (B) is its difficulty. The third parameter (C) is the lower asymptote parameter, which reflects the possibly non-zero chance of a correct response independent of ability. However, since the PIAAC test did not use any multiple-choice questions, parameter C was set at zero throughout, and the equation became a two-parameter logistic model, so to speak. After the parameters are calculated, each item is assigned a response probability of 67 (RP67), which represents the proficiency level that a respondent must have to complete the task with an 67% probability of success.

As noted previously, a respondent's proficiency was summarized on the three scales using the item parameters and the respondent's ability based on the IRT scaling models. The application differed somewhat from the norm in that in the PIAAC, each respondent was given a relatively small number of items so that the population's proficiency levels could be identified more effectively. However, because the data are not intended to estimate individual proficiency levels, more complicated analyses are required.

Plausible values methodology was used to estimate key population features consistently and to calculate others with the same degree of accuracy as standard IRT procedures. Essentially, this added dimension requires that the estimation of proficiency be based on a series of 10 plausible values for each of the two literacy domains. These 10 plausible values—pvlit1 through pvlit10 for the literacy scale, and pvnum1 through pvnum10 for the numeracy scale—were recoded into plausible levels with values from 0 to 5 reflecting the empirically determined progression of the information-processing skills and strategies required to perform increasingly complex tasks. The category 'Below Level 1' (coded 0 for these variables) is equivalent to scores in the range 0 to less than 176; Level 1, to scores of 176 to 225.9999; Level 2, to scores of 226to 275.9999; Level 3, to scores of 276to 325.9999; Level 4, to scores of 326 to 375.9999; and Level 5, to scores greater than or equal to 376. The variables are pvlitlev1 to pvlitlev10 for the literacy scale and pvnumlev1 to pvnumlev10 for the numeracy scale.

Because of a difference in the framework, the Problem Solving scale was constructed slightly differently. First, the same 10 plausible values ranging from 0 to 500 (pvpsl1 to pvpsl10) were created, but the definition of the levels was slightly different. For instance, this scale has only three proficiency levels, Level 1 being the lowest and Level 3 the highest, with an additional category, "below Level 1" (coded to 0 in these variables). To calculate simple point estimates for any of the 10 literacy domains, it is sufficient to use the population weight along with one of the 10 corresponding plausible values (chosen at random).

Note that taking only one of the plausible values produces only a valid point estimate, not a valid variance estimate. **All 10** plausible values and the 80 replicate weights must be used in order to correctly compute design-based variance estimates. Design-based variance estimates are discussed further in section 0 (Using Plausible Values and Replicate Weights in Calculating Sampling Errors).

11 Data Quality

The data quality from any survey can be evaluated by looking at two types of survey errors: sampling error and non-sampling error.

The estimates derived from this survey are based on a sample of individuals. Somewhat different figures might have been obtained if a complete census had been taken using the same questionnaire, interviewers, supervisors, processing methods, etc. as those actually used. The difference between the estimates obtained from the sample and the results from a complete count taken under similar conditions is called the sampling error of the estimate.

Errors, which are not related to sampling, may occur at almost every phase of a survey operation. Interviewers may misunderstand instructions, respondents may make errors in answering questions, the answers may be incorrectly entered on the questionnaire and errors may be introduced in the processing and tabulation of the data. These are all examples of non-sampling errors.

The following section covers sampling errors and non-sampling errors, but as well as data validation measures undertaken during collection.

11.1 Validation

The International Consortium regarded Quality Control (QC) as an integral component to the overall success of the Program for the International Assessment of Adult Competencies (PIAAC). Various guidelines were established to ensure that the data collected by participating countries were reliable and valid.

The guidelines stipulated that throughout collection PIAAC countries routinely conduct validations to verify that an interview was indeed conducted or attempted as reported by the interviewer. Countries were required to validate at least 10 percent of each interviewer's finalized work to ensure that the case was handled according to study procedures. Validation included completed cases and those finalized with other outcome codes, such as vacant or refusal. Cases were selected randomly.

In Canada, the Quality Control Validation (QCVAL) was done by a Computer Assisted Telephone Interview (CATI). The interview consisted of a series of questions about the respondents experience with the PIAAC survey, and the responses were then compared to the Main Study data to determine if:

- The data matched (month and year of birth; education; address; demographics on household members; etc)
- Procedures were followed (length of interview; composure of interviewer; interviewer using laptop; respondent completing assessment; interviewer helping respondent)
- The correct outcome code was assigned (correct vacant/ no contact/ absent/ seasonal dwelling etc).

If inconsistencies were discovered, the interviewer's entire completed caseload was then selected and subject to further validation in order to ascertain whether other cases were also compromised.

11.2 Sampling errors

Since it is an unavoidable fact that estimates from a sample survey are subject to sampling error, sound statistical practice calls for researchers to provide users with some indication of the magnitude of this sampling error. This section of the documentation outlines the measures of sampling error which Statistics Canada commonly uses and which it urges users producing estimates from this microdata file to use also.

The basis for measuring the potential size of sampling errors is the standard error of the estimates derived from survey results.

However, because of the large variety of estimates that can be produced from a survey, the standard error of an estimate is usually expressed relative to the estimate to which it pertains. This resulting measure, known as the coefficient of variation (C.V.) of an estimate, is obtained by dividing the standard error of the estimate by the estimate itself and is expressed as a percentage of the estimate.

For example, suppose that, based upon the survey results, one estimates that 16.6% of Canadians are at Level 1 with regard to literacy, and this estimate is found to have standard error of 0.013. Then the coefficient of variation of the estimate is calculated as:

$$\left(\frac{.013}{.166}\right) x\ 100\% \ = \ 7.8\%$$

11.2.1 CV release guidelines

One criterion that can used to determine whether survey estimates are publishable is the coefficient of variation (CV). The CV is the standard error of an estimate expressed as a percentage of that estimate.

Before releasing and/or publishing any estimate from the PIAAC, users should first determine the quality level of the estimate. The quality levels are acceptable, marginal and unacceptable. Data quality is affected by both sampling and non-sampling errors. However for release purposes, the quality level of an estimate will be determined only on the basis of sampling error and imputation error (when using plausible values) as reflected by the coefficient of variation as shown in table 11.1. Nonetheless users should be sure to read section 11 to be more fully aware of the quality characteristics of these data.

First, the unweighted number of respondents who contribute to the calculation of the estimate should be determined. If this number is less than 5, the result should not be published for confidentiality reason. For weighted estimates based on sample sizes of 5 or more, users should determine the coefficient of variation of the estimate and follow the guidelines below. These quality level guidelines should be applied to weighted rounded estimates. All estimates can be considered releasable.

However, those of marginal or unacceptable quality level must be accompanied by a warning to caution subsequent users.

Table 11.1 : Quality level guidelines

Quality level of estimate	**Guidelines**
1. Acceptable	Estimates have: a low coefficients of variation in the range 0.0% to 16.5%. No warning is required.
2. Marginal	Estimates have: a high coefficients of variation in the range 16.6% to 33.3%. Estimates should be flagged with the letter M (or some similar identifier). They should be accompanied by a warning to caution subsequent users about the high levels of error associated with the estimates.
3. Unacceptable	Estimates have: a very high coefficients of variation in excess of 33.3%. Statistics Canada recommends not to release estimates of unacceptable quality. However, if the user chooses to do so then estimates should be flagged with the letter U (or some similar identifier) and the following warning should accompany the estimates: “The user is advised that . . . (specify the data) . . . do not meet Statistics Canada’s quality standards for this statistical program. Conclusions based on these data will be unreliable, and most likely invalid. These data and any consequent findings should not be published. If the user chooses to publish these data or findings, then this disclaimer must be published with the data.”

11.2.2 Using plausible values and replicate weights in calculating sampling errors

The following section has been liberally copied from the documentation that accompanies the STATTOOL (SAS and SPSS) designed at the time to help users manipulate the ALL and IALSS data. While some details of the following section may be more particular to international comparisons of the type facilitated by the ALL Public Use Microdata file, the discussion that ensues will shed some light on the proper usages and practical limits of the PIAAC data as well. New tools have been developed by the international consortium and the OECD and available with documentation on the OECD website.

http://www.oecd.org/site/piaac/publicdataandanalysis.htm

11.2.2.1 Calculating point estimates

In this section, we will see how to use the sampling weights (SPFWT0) to obtain population estimates such as percentages (totals) and means.

All examples will be based on a fictional population with the following characteristics:

Using plausible values and replicate weights

There is little doubt that the PIAAC dataset is difficult to manipulate. The 10 Plausible values for the 3 domains, along with the 80 replicate weights make the procedures for accurate assessment of standard errors a convoluted affair.

In many instances, simplification of the process, particularly at the exploratory stage would greatly cut down on the processing time required to output the skill estimate analysis.

For this reason, it is recommended that preliminary research use only one of the Plausible values, rather than all ten. This is much more accurate than averaging the ten plausible values, since it allows for the weighted population distribution to accurately reflect the point estimate. The average of the PV would mask the testing error and, as the population under investigation gets smaller, the estimates will increasingly diverge from the true population distribution.

Of course, once the research is ready for publication, the replicate weights and 10 plausible values should be used to produce the final estimates with accurate standard errors. A full description of this procedure can be found beginning in section 11.2.2.

Type	Gender	Population distribution	Sample: Distribution (unweighted)	Sample: Distribution (weighted)
Rural	Male	40 %	30 %	38 %
	Female	60 %	70 %	62 %
	Total	20 %	50 %	19 %
Urban	Male	51 %	45 %	50 %
	Female	49 %	55 %	50 %
	Total	80 %	50 %	81 %
Total	Male	48.8 %	37.5 %	47.7 %
	Female	51.2 %	62.5 %	52.3 %
	Total	100 %	100 %	100 %

From this table, it seems that the male participation was lower than the female participation in both rural and urban areas. Even though nearly 49% of the population is made of males, we have only 37.5% males in the sample. This trend can be observed in both areas. It seems also that the rural area was over allocated with 50% of the sample coming from that area compared to only 20% in the population.

However, once the sampling weights are used, the percentages are quite comparable. How are they calculated?

Percentages (Totals)
The weighted percentage of the males living in rural areas was calculated as follow:

$$Weighted\ \% = \frac{\sum_{i=1}^{rural,male} SPFWT0_i}{\sum_{i=1}^{rural} SPFWT0_i} = 38\%$$

where i identifies individual i. The numerator is an estimate of the total population of males living in rural areas while the denominator is an estimate of the total population living in the rural areas.

The unweighted percentage was calculated as follow:

$$Unweighted\ \% = \frac{\sum_{i=1}^{rural,male} 1_i}{\sum_{i=1}^{rural} 1_i} = \frac{n_{rural,male}}{n_{rural}} = 30\%$$

where $n_{rural,male}$ is the total number of males living in rural areas found in the sample and n_{rural} is the total number of people living in rural areas found in the sample.

In the latter, each sampled individual accounts for one while in the weighted version, each sampled unit was given a weight in order to properly and proportionally represent the subgroups in the sample (note that the weighted percentage is a ratio of estimated weighted totals).

Means

For this fictional example, let's say that we also have the average score based on variable PVLIT1 as illustrated by the following table:

Type	Gender	Population distribution	Sample			
			Distribution (unweighted)	Distribution (weighted)	Avg. Pvlit1 (unweighted)	Avg. Pvlit1 (weighted)
Rural	Male	40 %	30 %	38 %	260	260.1
	Female	60 %	70 %	62 %	290	289.8
	Total	20 %	50 %	19 %	281	278.5
Urban	Male	51 %	45 %	50 %	320	319.7
	Female	49 %	55 %	50 %	330	330.1
	Total	80 %	50 %	81 %	325.5	324.9
Total	Male	48.8 %	37.5 %	47.7 %	296.0	310.7
	Female	51.2 %	62.5 %	52.3 %	307.6	321.0
	Total	100 %	100 %	100 %	303.3	316.1

Here again we see that the weighted means are quite close to the unweighted means as long as one controls by area type. This is not true for the last 3 lines of the table. Let's try to see why. The weighted mean of the males living in rural areas was calculated as follow:

$$Weighted\ mean = \frac{\sum_{i=1}^{rural,male} SPFWT0_i * PVLIT1_i}{\sum_{i=1}^{rural,male} SPFWT0_i} = 260.1$$

where i identifies individual i. The numerator is an estimate of the total score for all males living in rural areas while the denominator is an estimate of the total male population living in rural areas.

The unweighted mean was calculated as follow:

$$Unweighted\ mean = \frac{\sum_{i=1}^{rural,male} PVLIT1_i}{\sum_{i=1}^{rural,male} 1_i} = \frac{\sum_{i=1}^{rural,male} PVLIT1_i}{n_{rural,male}} = 260$$

The unweighted and weighted results will be similar whenever values of PVLIT1 don't vary much from one individual to the other and/or values of SPFWT0 behave the same way. This statement doesn't hold for the last three lines of the table. The weighted mean for male is obtained by solving the following equation:

$$Weighted\ \ mean = \frac{\sum_{i=1}^{male} SPFWT0_i * PVLIT1_i}{\sum_{i=1}^{male} SPFWT0_i}$$

$$= \frac{(38\% * 19\% * 260.1) + (50\% * 81\% * 319.7)}{47.7\%} = 310.7$$

While the unweighted mean is given by:

$$Unweighted\ \ mean = \frac{\sum_{i=1}^{male} PVLIT1_i}{\sum_{i=1}^{male} 1_i}$$

$$= \frac{(30\% * 50\% * 260) + (45\% * 50\% * 320)}{37.5\%} = 296.0$$

In the latter, each sampled individual accounts for one while in the weighted version, each sampled unit was given a weight in order to properly and proportionally represent the subgroups in the sample. For example, the 30%X50%=15% of males living in rural areas found in the sample was adjusted by the weights to account for 38%X19%=7.22% of the entire sample which is a much better reflect of what is found in the whole population (Note that the true population proportion of males living in rural areas is 40%X20%=8%).

In conclusion, any statistics computed form sample data should always be done using the sampling weights.

11.2.2.2 Alternative sampling weights

As we saw earlier, the sum of the sampling weights under SPFWT0 within a sample provides an estimate of the size of the population. Although this is a commonly used sampling weight, it sometimes adds to a very large number, and to different numbers from country to country. This is not always desirable. For example, if you want to compute a weighted estimate of the mean achievement in the population across all countries (or sub-populations within a country), using the variable SPFWT0 as your weight variable will lead each country to contribute proportionally to its population size, with the larger countries counting more than small countries. In general, SPFWT0 is not the weight of choice for cross-country analyses. Another consequence of using SPFWT0 is the tendency to inflate results in significance tests when computer softwares are unable to deal correctly with weighted data. We will now see two possible versions of individual sampling weights that address these issues in particular. These versions take advantage on the fact that the same population estimates for means and proportions is obtained whenever you use a weight variable proportional to the population weight (SPFWT0).

Sum to Constant Sampling Weight (CONSTWT)

It is possible to modify the population weight SPFWT0 such that all countries would contribute the same in a cross-country mean or proportion. This is given by:

$$CONSTWT_{g,i} = SPFWT0_{g,i} * \left[\frac{100}{\sum_{i=1}^{g} SPFWT0_i} \right]$$

for each individual in the group of interest g. The transformation of the weights will be different within each country, but in the end the sum of the variable CONSTWT within each country will be 100. The variable CONSTWT, within each country, is proportional to SPFWT0 multiplied by the ratio of 100 divided by the sum of weights over all individuals in the group of interest. These weights can be used when international estimates are sought and you want to have each country contribute the same amount to the international estimate, regardless of the size of the group of interest in the country (see table below).

Group of interest	Country	Population count (rural)	Population estimates	
			Mean PVLIT1 (SPFWT0)	Mean PVLIT1 (CONSTWT)
Rural	A	3,700,000	290	290
	B	37,000,000	260	260
	C	7,000,000	300	300
Overall			268	283

Sum to Sample Size Sampling Weight (SMPLWT)

It is possible to modify the population weight SPFWT0 when you want the actual sample size to be used in performing significance tests (within each country). This is given by:

$$SMPLWT_{g,i} = SPFWT0_{g,i} * \left[\frac{n_g}{\sum_{i=1}^{g} SPFWT0_i} \right]$$

for each individual in the group of interest g where n_g is the actual sample size in group g. The transformation of the weights will be different within each country, but in the end the sum of the variable SMPLWT within each country will add up to the sample size in group g. The variable SMPLWT, within each country, is proportional to SPFWT0 multiplied by the ratio of the sample size (n_g) divided by the sum of weights over all individuals in the group of interest. Although some statistical computer software packages allow you to use the sample size as the divisor in the computation of standard errors, others will use the sum of the weights, and this results in severely deflated standard errors for the statistics if SPFWT0 is used as the weighting variable. When performing analyses using such software, it is recommended to use a weighting variable such as SMPLWT as the weight variable. Because of the clustering effect in most country samples, it may also be desirable to apply a correction factor such as a design effect to the SMPLWT variable.

11.2.2.3 Using the plausible values to compute point estimates

To achieve its goal of broad coverage of assessment of adult competencies, the PIAAC assessment included a range of items arranged into assessment blocks. Each individual participating in the assessment completed a subset of items keeping individual response burden to a minimum. PIAAC used a matrix-sampling design to assign assessment blocks to individuals so that a comprehensive picture of the literacy, numeracy and PS-TRE achievement in each country could be assembled from the components completed by each individual. PEICA relied on Item Response Theory (IRT) scaling to combine the individual responses to provide accurate estimates of the three domains of achievement in the population in each country. The PIAAC IRT scaling also uses multiple imputation or "plausible values" methodology to obtain proficiency scores for all individuals, even though each individual responded to only a part of the assessment item pool.

Most cognitive skills testing is concerned with accurately assessing the performance of individual respondents for the purposes of diagnosis, selection or placement. The accuracy of these measurements can be improved by increasing the number of items given to the individual. For the distribution of proficiencies in large population, however, more efficient estimates can be obtained from matrix-sampling design. These designs solicit few responses from each sampled respondent while maintaining a wide range of content representation when responses are aggregated across all respondents. With this approach, however, the advantage of estimating population characteristics is more efficiently offset by the inability to make precise statements about individuals, with the result that aggregations of individual scores can lead to seriously biased estimates of population characteristics.

Plausible values methodology was developed as a way to address this issue by using all available data to estimate directly the characteristics of populations and sub-populations, and then generating multiple imputed scores (called plausible values) from these distributions, which can be used in analyses with standard statistical software. A detailed review of plausible values methodology is given by Mislevy (1991). The main things to retain from this are:

a) Whenever you want to compute statistics involving scores (like PVLIT, PVNUM, etc) you don't have one score value but ten score values assigned to each individuals. Each set of plausible values is equally well-designed to estimate population parameters;

b) These statistics based on scores should always be computed at population or subpopulation levels. They should never be used to do inference at individual level.

11.2.2.4 Working with plausible values

Example1: Estimated median for the variable PVLIT in Country A.

For each individual, we don't have one but 10 scores to deal with as illustrated in the next table

Country A	PVLIT1	PVLIT2	PVLIT3	PVLIT4	PVLIT5	PVLIT6	PVLIT7	PVLIT8	PVLIT9	PVLIT10
Individual 1	222	275	300	245	254	242	265	290	255	234
Individual 2	289	310	212	250	268	293	306	211	252	267
…	…	…	…	…	…	…	…	…	…	…
Individual n	285	275	243	321	312	275	235	293	331	302
Median	285	281	283	279	289	283	287	285	276	281

In order to estimate the overall median for PVLIT in country A we first have to estimate the median based on the first set of plausible values found under variable PVLIT1. We then repeat this first step using PVLIT2 through PVLIT10 to get a total of ten equally good estimates of the median for that country. Since they are all equally good, the next step to do to obtain a single estimate of the median is to average out these ten estimates. We get:

Overall Median = (285 + 281 + 283 + 279 + 289 + 283 + 287 + 285 + 276 + 281) / 10 = 282.9

Note that you should not average out scores at the individual level. For example, (222 + 275 + 300 + 245 + 254 + 242 + 265 + 290 + 255 +234) / 10 = 258.2 is not a good estimate of the variable PVLIT for individual 1. This is true for any of the PVLITn variables for a given individual. Found score variables should always be interpreted in populations or subpopulations context.

The variables PVLIT1 through PVLIT10 are the raw ability scores for the literacy domain. When these scores are re-grouped into levels, they yield ten level variables called PVLITLEV1 through PVLITLEV10.

Example 2: Estimated logistic regression parameter coefficients (Levels 2 versus level 3 of the dependent variables PVLITLEV1 to PVLITLEV10).

From the values found in the previous table this would give:

Country A	LEV1	LEV2	LEV3	LEV4	LEV5	LEV6	LEV7	LEV8	LEV9	LEV10
Individual 1	2	3	3	2	2	2	2	3	2	2
Individual 2	3	3	1	2	2	3	3	1	2	2
…	…	…	…	…	…	…	…	…	…	…
Individual n	3	2	2	3	3	3	3	3	3	3

Note: for presentation purposes, the PVLITLEV1 to PVLITLEV10 are changed in the table to LEV1 to LEV10

The first thing we note is that a given individual may be found in different proficiency levels depending on which set of plausible values we are looking at. This does not invalidate the methodology used however. As explained before, in order to get the logistic regression parameter coefficients at the country level, we first have to calculate these parameters based on the first set of plausible values. We then repeat this step 9 times using PVLITLEV2 through PVLITLEV10 to get 9 additional sets of estimated parameters as illustrated in the next table:

Country A	LEV1	LEV2	LEV3	LEV4	LEV5	LEV6	LEV7	LEV8	LEV9	LEV10
Intercept	0.124	0.129	0.122	0.125	0.120	0.121	0.125	0.132	0.124	0.126
Beta 1	1.051	1.059	1.049	1.055	1.060	1.043	1.079	1.044	1.055	1.030
Beta 2	0.584	0.591	0.545	0.499	0.645	0.684	0.541	0.585	0.559	0.642
Beta 3	3.222	4.123	3.012	3.542	3.201	3.223	4.193	3.011	3.546	3.251

Note : PVLITLEV1 to PVLITLEV10 were replace by LEV1 to LEV10 in the table for presentation purpose

Since these sets of estimated parameter coefficients are all equally good, the next step to do to obtain a single set of estimates is to average out these ten results. We get:

Overall Intercept =
(0.124 + 0.129 + 0.122 + 0.125 + 0.120 + 0.121 + 0.125 + 0.132 + 0.124 + 0.126) / 10 = 0.125

Overall Beta 1 =
(1.051 + 1.059 + 1.049 + 1.055 + 1.060 + 1.043 + 1.079 + 1.044 + 1.055 + 1.030) / 10 = 1.052

Overall Beta 2 =
(0.584 + 0.591 + 0.545 + 0.499 + 0.645 + 0.684 + 0.541 + 0.585 + 0.559 + 0.642) / 10 = 0.588

Overall Beta 3 =
(3.222 + 4.123 + 3.012 + 3.542 + 3.201 + 3.223 + 4.193 + 3.011 + 3.546 + 3.251) / 10 = 3.432

11.2.3 Estimating error variance in PIAAC

The PIAAC methodology is a tried and tested method of quantifying skills. As such, a large body of documentation already exists regarding the appropriate ways of estimating variance in such studies. A large part of the text found in this section was borrowed from the IEA PIRLS 2001 User's Guide and adapted to fit the PIAAC context.

11.2.3.1 Overview

When analyzing data from complex designs such as PIAAC, it is important to compute correct error variance estimates for the statistics of interest. In PIAAC this error variance can come from two sources: the sampling process (always present) and the imputation process (whenever the statistics of interest involve proficiency scores). This section describes the methods used to estimate these error variance components.

11.2.3.2 Estimating sampling variance

When data are collected as part of a complex sample survey, analytically there is often no easy way to produce unbiased or design-consistent estimates of variance. A class of techniques called replication methods provides a way to estimate variance for the type of complex sample designs such as those used in PIAAC.

The basic idea behind replication is to select subsamples repeatedly from the whole sample, calculate the statistic of interest for each subsample, and then use these subsamples or replicate statistics to estimate the variance of the full-sample statistic. Different ways of creating subsamples from the full-sample result in different replication methods. The subsamples are called replicates and the statistics calculated from these replicates are called replicate estimates.

One such method is the jackknife repeated replication (JRR) technique (Wolter, 1985). In Canada for PIAAC, the full sample was randomly split into 80 subsets of equal or nearly equal size, with each subset resembling the full sample. Replicates are formed by deleting one subset at a time and multiplying the weights for the other subsets by

$$\frac{80}{79}$$

In this manner, 80 replicates are created. This method is also known as the JK1 method. The weights associated with each replicate can be found under variables SPFWT1 through SPFWT80. These weights should only be used to calculate the sampling variance. Point estimates should be calculated as described in the previous section.

11.2.3.3 Computing sampling variance using the JK1 method

The basic idea here is to calculate the estimate of interest from the full sample using the weight variable SPFWT0 (or SMPLWT or CONSTWT) as well as each replicate (using the variables SPFWT1 through SPFWT80). The variation between the replicate estimates and the full-sample estimate is then used to estimate the sampling variance for the full sample. Suppose that $\hat{\theta}$ is the full-sample estimate of some population parameter θ. The sampling variance estimator $\mathrm{var}(\hat{\theta})$ is given by:

$$\mathrm{var}_{smpl}(\hat{\theta}) = \frac{79\sum_{g=1}^{80}(\hat{\theta}_{(g)} - \hat{\theta})^2}{80}$$

where $\hat{\theta}_{(g)}$ is the estimate of θ based on the observations included in the g-th replicate.

Example 1 : Average personal income in country A

The average personal income is given by the following expression:

$$Mean\ \ personal\ income = \frac{\sum_{i=1}^{CountryA} SPFWT0_i * D43_i}{\sum_{i=1}^{CountryA} SPFWT0_i} = 27{,}601$$

In order to compute the sampling variance, we have to calculate the following expression 80 times, each one based on the appropriate replicate weight.

$$Mean\ personal\ income_{(1)} = \frac{\sum_{i=1}^{CountryA} SPFWT1_i * D43_i}{\sum_{i=1}^{CountryA} SPFWT1_i} = 26{,}983$$

$$Mean\ personal\ income_{(2)} = \frac{\sum_{i=1}^{CountryA} SPFWT2_i * D43_i}{\sum_{i=1}^{CountryA} SPFWT2_i} = 26{,}146$$

$$\bullet\bullet\bullet$$

$$Mean\ personal\ income_{(80)} = \frac{\sum_{i=1}^{CountryA} SPFWT80_i * D43_i}{\sum_{i=1}^{CountryA} SPFWT80_i} = 28{,}965$$

We then simply apply the variance formula given earlier. This gives:

$$\text{var}_{smpl}(\hat{\theta}) = \frac{79\sum_{g=1}^{80}(\hat{\theta}_{(g)} - \hat{\theta})^2}{80} = \frac{79}{80}\left[(26{,}983 - 27{,}601)^2 + (26{,}146 - 27{,}601)^2 + \cdots + (28{,}965 - 27{,}601)^2\right]$$

Finally, the statistic $(\hat{\theta} - \theta) / \text{var}(\hat{\theta})^{1/2}$ is approximately t-distributed with 79 degrees of freedom.

11.2.3.4 Estimating imputation variance

Whenever the statistics of interest involve proficiency scores, there is a need for estimating the imputation variance. As mentioned in previous section, ten scores are generated for the same test for individuals participating in PIAAC. These different scores are referred to as plausible values (PVs). In PIAAC, each individual was presented with a set of items. The full collection of items in PIAAC covers the concepts to be tested, but an individual respondent did not answer questions from all the items. Using a type of balanced assignment of items of respondents, the full battery of questions was covered when respondents are aggregated. For a group of similar respondents, a Bayesian posterior distribution of scores was estimated. The plausible values for each respondent are realizations from the posterior distribution. These scores are not meaningful for an individual respondent, but when combined can be used to estimate population averages and other population quantities.

11.2.3.5 Computing imputation variance

The general procedure for estimating the imputation variance using plausible values is as follows:

a) First estimate the statistic of interest θ, each time using a different set of plausible values (M) and the variable SPFWT0 (or SMPLWT or CONSTWT). Let's call these 10 estimates, $\hat{\theta}_1$ to $\hat{\theta}_{10}$. The statistic of interest can be anything estimable from the sample data, such as mean, the difference between means, percentiles, regression parameter coefficients, etc.
b) Then estimate the overall estimate by averaging out the $\hat{\theta}_m$ where m=1, 2,…, 10. Let's call this estimate $\hat{\theta}$.
c) The imputation variance is computed as:

$$Var_{imp}(\hat{\theta}) = \left[1 + \frac{1}{10}\right] \times \sum_{m=1}^{10} \frac{\left(\hat{\theta}_m - \hat{\theta}\right)^2}{9}$$

11.2.3.6 Estimating the overall error variance

The overall error variance is then computed as follows:

$$Var(\hat{\theta}) = \sum_{m=1}^{10} \frac{Var_{smpl}(\hat{\theta}_m)}{10} + Var_{imp}(\hat{\theta})$$

When the statistics of interest do not involve any proficiency scores, the overall error variance formula simply becomes:

$$Var(\hat{\theta}) = Var_{smpl}(\hat{\theta})$$

11.2.3.7 Degrees of freedom

$(\hat{\theta} - \theta) / Var(\hat{\theta})^{1/2}$ is approximately t-distributed, with degrees of freedom (Jonhson & Rust,1993) given by:

$$\nu = \frac{1}{\frac{f_m^2}{9} + \frac{(1 - f_m)^2}{79}}$$

where f_m is given by:

$$f_m = \frac{Var_{imp}(\hat{\theta})}{Var(\hat{\theta})}.$$

In practice, the number of degrees of freedom is set to 79 (this will work well when f is relatively small, say less than 30%).

11.2.3.8 Making comparisons

We will now see how to compute the correct error variance when comparing survey estimates between countries, to the international estimates, and within countries. In order to simplify the text, the estimated mean achievement for the variable PVLIT will be considered only. It should be straightforward to generalize this section to any type of survey estimate.

11.2.3.9 Between countries

The error variance when comparing the estimated mean achievement for PVLIT between country A and B is given by:

$$Var(\hat{\theta}) = Var_{CountryA}(\hat{\theta}) + Var_{CountryB}(\hat{\theta})$$

For example, say that the estimated mean achievement for country A is 290 with an error variance of 25 and for country B, the estimated mean achievement is 307 with an error variance of 30.25. The difference between these two countries is then 307-290 = 17. The question is : is this difference of 17 points the result of error due to sampling only part of the population combined with the fact that only part of the items were administered? To find the answer to that question, we first have to compute the following statistic (known as the Wald statistic):

$$(\hat{\theta}) / Var(\hat{\theta})^{1/2} = 17/(25 + 30.25)^{1/2} = 2.287$$

When this value is compared to the critical 95% value from a t distribution with 79 degrees of freedom (1.99), we conclude that there is enough evidence to state that these two countries don't have the same estimated mean achievement.

Note that this approach is also valid when comparing the IALSS results to the PIAAC results.

11.2.3.10 To the international estimates

An important published statistics shows your country mean achievement compared to the international mean. The error variance when doing so is given by:

$$Var(\hat{\theta}) = \frac{(N-1)^2 Var_{smpl\,A}(\hat{\theta}_A) + \sum_{k=1, k \neq A}^{N} Var_{smpl\,k}(\hat{\theta}_k)}{N^2} + Var_{imp}(\hat{\theta}_A - \hat{\theta}_{\text{int}})$$

where N is the number of countries used to compute the international mean, $\hat{\theta}_m$ stands for the estimated mean for country m, and $\hat{\theta}_{\text{int}}$ stands for the estimated international mean.

For example, let's consider the following tables:

	Mean achievement	Sampling variance
Country A	290	20
Country B	300	22
Country C	286	18
Country D	324	22
International	300	

Mean achievement	PVLIT1	PVLIT2	PVLIT3	PVLIT4	PVLIT5	PVLIT6	PVLIT7	PVLIT8	PVLIT9	PVLIT10
Country A	288	292	292	288	290	289	291	293	288	289
International	301	300	299	302	298	300	302	298	301	299
$\hat{\theta}_A - \hat{\theta}_{\text{int}}$	-13	-8	-7	-14	-8	-11	-11	-5	-13	-10

Here we have that the estimated mean achievement for country A is 290 with a sampling variance of 20 (imputation variance of 4) and the estimated international mean achievement based on 4 countries is 300. The difference between country A result and the international mean is then 290-300 = -10. The question is : is this difference of 10 points the result of error due to sampling only part of the population combined with the fact that only part of the items were administered? To find the answer to that question, we first have to compute the following statistic:

$$Var(\hat{\theta}) = \frac{(3)^2 20 + (22+18+22)}{4^2} +$$

$$1.1 * \frac{\left[(13-10)^2 + (8-10)^2 + (7-10)^2 + (14-10)^2 + (8-10)^2 + (11-10)^2 + (11-10)^2 + (5-10)^2 + (13-10)^2 + (10-10)^2\right]}{9}$$

$$= 15.125 + 9.533 = 24.658$$

We then compute the Wald statistics:

$$(\hat{\theta}) / Var(\hat{\theta})^{1/2} = 10/(24.658)^{1/2} = 2.014$$

When this value is compared to the critical 95% value from a t distribution with 79 degrees of freedom (1.99), we conclude that there is enough evidence to state that country A is different from the estimated international mean achievement.

11.2.3.11 Within countries

Most of the times when comparing subgroups within countries, there is no direct formula for computing the overall error variance like we had in the previous two sections. The main reason for this is that, the samples for the different subgroups are not typically treated as independent for the purpose of statistical tests. Accordingly, a jackknife procedure applicable to correlated samples for estimating the sampling variance of the difference between subgroups should be applied. This involves computing the difference between subgroups once for each of the 80 replicate samples, and then repeat nine more times, once for each set of plausible values as described earlier (see combining sampling and imputation variances).

However, linear regression models can be easily used to compute these differences. Here's how to compute the difference between men and women for the variable PVLIT:

a) Create dummy variables for the subgroups; let's call MAN the variable that will take value 1 if the respondent is a man and 0 otherwise, and WOMAN the variable that will take value 1 if the respondent is a woman and 0 otherwise. You create as many dummy variables as there are subgroups.
b) Using SPFWT0, run a linear regression model with PVLIT1 as the dependent variable and either MAN or WOMAN as the independent variable (When there are more k dummy variables with k greater than 2, choose k-1 dummy variables as independent variables). The dummy variable left out will become the reference subgroup.
c) Using the replicate weights, repeat step b), 80 times.
d) Repeat step b) and c) 9 times, once for each set of plausible values
e) Combine information from step b), c), and d) to compute the overall point estimate, the error variance, and the Wald statistics.

For example, say that after creating the dummy variables you get from step b), the following result:

Mean PVLIT1 = 270 + 18*WOMAN (weighted by SPFWT0)

This simplifies to Mean PVLIT1 = 270 when respondents are males and Mean PVLIT1 = 288 when respondents are females. This means that the coefficient in front of the variable WOMAN in the regression model is the difference between women and men while the intercept (270) is the mean achievement for the reference level, men in this case.

From step c), you get:

Mean PVLIT1 = 271 + 17*WOMAN (weighted by SPFWT1)
Mean PVLIT1 = 269 + 19*WOMAN (weighted by SPFWT2)
Mean PVLIT1 = 273 + 14*WOMAN (weighted by SPFWT3)
…

Mean PVLIT1 = 268 + 21*WOMAN (weighted by SPFWT80)

The sampling variance of the difference between women and men can now be calculated as:

$$\text{var}_{smpl}(\hat{\theta}) = \frac{79\sum_{g=1}^{80}(\hat{\theta}_{(g)} - \hat{\theta})^2}{80} = \frac{79}{80}\left[(17-18)^2 + (19-18)^2 + (14-18)^2 + \cdots + (21-18)^2\right] = 19.575$$

From step d), you get:

Mean PVLIT2 = 271 + 20*WOMAN (weighted by SPFWT0), $\text{var}_{smpl}(\hat{\theta}) = 19.75$

Mean PVLIT3 = 269 + 19*WOMAN (weighted by SPFWT0), $\text{var}_{smpl}(\hat{\theta}) = 20.125$

Mean PVLIT4 = 273 + 17*WOMAN (weighted by SPFWT0), $\text{var}_{smpl}(\hat{\theta}) = 19.425$

Mean PVLIT5 = 268 + 16*WOMAN (weighted by SPFWT0), $\text{var}_{smpl}(\hat{\theta}) = 19.675$

Mean PVLIT6 = 272 + 21*WOMAN (weighted by SPFWT0), $\text{var}_{smpl}(\hat{\theta}) = 18.95$

Mean PVLIT7 = 268 + 18*WOMAN (weighted by SPFWT0), $\text{var}_{smpl}(\hat{\theta}) = 19.5$

Mean PVLIT8 = 274 + 17*WOMAN (weighted by SPFWT0), $\text{var}_{smpl}(\hat{\theta}) = 19.65$

Mean PVLIT9 = 269 + 17*WOMAN (weighted by SPFWT0), $\text{var}_{smpl}(\hat{\theta}) = 19.325$

Mean PVLIT10 = 271 + 17*WOMAN (weighted by SPFWT0), $\text{var}_{smpl}(\hat{\theta}) = 19.775$

The overall point estimate for the difference can now be computed by averaging out the results over PVLIT1 to PVLIT10. This gives: (18+20+19+17+16+21+18+17+17+17) / 10 = 18. The sampling variance is given by:

$$Var(\hat{\theta}) = \sum_{m=1}^{10}\frac{Var_{smpl}(\hat{\theta}_m)}{10} = \frac{19.575+19.75+20.125+19.425+19.675+18.95+19.5+19.65+19.325+19.775}{10}$$
$$= 19.575$$

The imputation variance is given by:

$$Var_{imp}(\hat{\theta}) = \left[1 + \frac{1}{10}\right] \times \sum_{m=1}^{10}\frac{\left(\hat{\theta}_m - \hat{\theta}\right)^2}{9}$$

$$= 1.1 * \frac{\left[(18-18)^2 + (20-18)^2 + (19-18)^2 + (17-18)^2 + (16-18)^2 + (21-18)^2 + (18-18)^2 + (17-18)^2 + (17-18)^2 + (17-18)^2\right]}{9}$$

$$= 2.689$$

The Wald statistics becomes

$$(\hat{\theta})/Var(\hat{\theta})^{1/2} = 18/(19.575 + 2.689)^{1/2} = 3.815$$

When this value is compared to the critical 95% value from a t distribution with 79 degrees of freedom (1.99), we conclude that there is enough evidence to state that men mean achievement is different than women mean achievement within country A.

This last section was to provide a broad overview on how to use correctly the PIAAC data. However, as mentioned previously, tools and macro have been developed by the international consortium and the OECD and can be downloaded on the OECD site. User manual for these tools are also available on the site.

11.3 Non-sampling errors

Over a large number of observations, randomly occurring non-sampling errors will have little effect on estimates derived from the survey. However, errors occurring systematically will contribute to biases in the survey estimates. Considerable time and effort was made to reduce non-sampling errors in the survey. Quality assurance measures were implemented at each step of the data collection and processing cycle to monitor the quality of the data. These measures included the use of highly skilled interviewers, extensive training of interviewers with respect to the survey procedures and questionnaire, observation of interviewers to detect problems of questionnaire design or misunderstanding of instructions, procedures to ensure that data capture errors were minimized and coding and edit quality checks to verify the processing logic.

Despite these efforts, non-sampling error is bound to exist in every survey. The following text outlines the most likely sources of this error and its impact on the PIAAC survey.

11.3.1 Sampling frame

The use of the 2011 Census insured that the PIAAC frame was as inclusive as possible and that any exclusions could be effectively taken into consideration into the overall survey design.

11.3.2 Non-response

A major source of non-sampling errors in surveys is the effect of non-response on the survey results. The extent of non-response varies from partial non-response (failure to answer just one or some questions) to total non-response.

Total non-response occurred when the interviewer was either unable to contact the respondent, no member of the household was able to provide the information, or the respondent refused to participate in the survey. The national non-response rate for the PIAAC was around 38%. Analysis of the PIAAC non-respondents characteristics tend to say that these people are concentrated in given groups (which means that the non-response does not seem to have been random). Non-response weighting adjustments were performed to compensate for total non-response. These adjustments were designed to reduce the non-response bias as much as possible, using, among others, variables that were linked to the response probability.

Partial non-response to the survey occurred, in most cases, when the respondent did not understand or misinterpreted a question, refused to answer a question, or could not recall the requested information. Generally, the extent of partial non-response was small in the PIAAC.

11.3.3 Response error

A number of other potential sources of non-sampling error that are unique to the PIAAC deserve comment. Firstly, some of the respondents may have found the test portion of the study intimidating and this may have had a negative effect on their performance. Unlike “usual” surveys, the PIAAC test items have “right” and “wrong” answers. Also, for many respondents this would have been their first exposure to a “test” environment in a considerable number of years. Further, although interviewers did not enforce a time limit for answering questions, the reality of having someone watching and waiting may have, in fact, imposed an unintentional time pressure. It is recognized, therefore that even though items were chosen to closely reflect everyday tasks, the test responses might not fully reveal the literacy capabilities of respondents due to the testing environment. Further, although the test nature of the study called for respondents to perform the activities completely independently of others, situations in the real world often enable persons to sort through printed materials with family, friends and associates. It could be therefore, that the skills measured by the survey do not reflect the full range of some respondents’ abilities in a more natural setting.

11.3.4 Scoring

Another potential source of non-sampling error for the PIAAC relates to the scoring of the test items, particularly those that were scored on a scale (e.g. items that required respondents to write). Special efforts such as centralizing the scoring and sample verification were made to minimize the extent of scoring errors. And as mentioned previously, a large proportion of the scoring was done by the computer; it increases importantly the quality of the scoring.

12 Coefficients of variance tables

12.1 Release cut-off's for the Programme for the International Assessment of Adult Competencies, 2012.

The following table provides an indication of the precision of population estimates as it shows the release cut-offs associated with each of the three quality levels presented in Section 11.2.1. These cut-offs are derived from the coefficient of variation (CV) tables discussed in Section 12.9.

For example, the table shows that the quality of a weighted estimate of 10,000 people possessing a given characteristic in the Atlantic Provinces is marginal.

Note that these cut-offs apply to estimates of population totals only. To estimate ratios, users should not use the numerator value (nor the denominator) in order to find the corresponding quality level. Rule 4 in Section 12.3 and Example 4 in Section 12.4 explain the correct procedure to be used for ratios.

Region	**Acceptable CV** **0.0% to 16.5%**		**Marginal CV** **16.6% to 33.3%**			**Unacceptable CV** **> 33.3%**	
Atlantic Provinces	17,000	& over	17,000	to <	4,000	under	4,000
Quebec	44,500	& over	44,500	to <	11,000	under	11,000
Ontario	135,000	& over	135,000	to <	33,500	under	33,500
Western Provinces	106,500	& over	106,500	to <	26,500	under	26,500
Northern Territories	4,000	& over	4,000	to <	1,000	under	1,000
Canada	**98,000**	**& over**	**98,000**	**to <**	**24,000**	**under**	**24,000**

12.2 Approximate sampling variability tables

In order to supply coefficients of variation (CV) which would be applicable to a wide variety of categorical estimates produced from this microdata file and which could be readily accessed by the user, a set of Approximate Sampling Variability Tables has been produced. These CV tables allow the user to obtain an approximate coefficient of variation based on the size of the estimate calculated from the survey data.

The coefficients of variation are derived using the variance formula for simple random sampling and incorporating a factor which reflects the multi-stage, clustered nature of the sample design. This factor, known as the design effect, was determined by first calculating design effects for a wide range of characteristics and then choosing from among these a conservative value (usually the 75^{th}

percentile) to be used in the CV tables which would then apply to the entire set of characteristics.

The table below shows the conservative value of the design effects as well as sample sizes and population counts by province which were used to produce the Approximate Sampling Variability Tables for the Programme for the International Assessment of Adult Competencies (PIAAC).

Region	**Design effect**	**Sample size***	**Population**
Atlantic Provinces	1.67	5,641	1,569,645
Quebec	1.34	5,875	5,371,975
Ontario	2.16	5,236	9,059,605
Western Provinces	3.23	7,803	7,114,465
Northern Territories	3.98	2,517	76,843
Canada	3.13	27,072	23,182,232

* : only individuals take were taken into effect in the score calculation are considered here.

All coefficients of variation in the Approximate Sampling Variability Tables are approximate and, therefore, unofficial. Estimates of actual variance for specific variables may be obtained from Statistics Canada on a cost-recovery basis or by using the 80 jackknife replicate weights, as described in Section 11.2. Since the approximate CV tends to be conservative, the use of actual variance estimates may cause the estimate to be switched from one quality level to another. For instance, a *marginal* estimate could become *acceptable* based on the exact CV calculation.

Note: The design effects of estimates for Francophones in the Western provinces are particularly high. Approximate CVs for this subpopulation should be multiplied by a factor of 1.75.

12.3 How to use the coefficient of variation tables for categorical estimates

The following rules should enable the user to determine the approximate coefficients of variation from the Approximate Sampling Variability Tables for estimates of the number, proportion or percentage of the surveyed population possessing a certain characteristic and for ratios and differences between such estimates.

Rule 1: Estimates of numbers of persons possessing a characteristic (Aggregates)
The coefficient of variation depends only on the size of the estimate itself. On the Approximate Sampling Variability Table for the appropriate geographic area, locate the estimated number in the left-most column of the table (headed "Numerator of Percentage") and follow the asterisks (if any) across to the first figure encountered. This figure is the approximate coefficient of variation.

Rule 2 : Estimates of proportions or percentages of persons possessing a characteristic
The coefficient of variation of an estimated proportion or percentage depends on both the size of the proportion or percentage and the size of the total upon which the proportion or percentage is based. Estimated proportions or percentages are relatively more reliable than the corresponding estimates of the numerator of the proportion or percentage, when the proportion or percentage is based upon a

sub-group of the population. For example, the proportion of females aged 16 to 65 who have completed post-secondary education is more reliable than the estimated number of females aged 16 to 65 who have completed post-secondary education. (Note that in the tables the coefficients of variation decline in value reading from left to right).

When the proportion or percentage is based upon the total population of the geographic area covered by the table, the CV of the proportion or percentage is the same as the CV of the numerator of the proportion or percentage. In this case, Rule 1 can be used.

When the proportion or percentage is based upon a subset of the total population (e.g. those in a particular sex or age group), reference should be made to the proportion or percentage (across the top of the table) and to the numerator of the proportion or percentage (down the left side of the table). The intersection of the appropriate row and column gives the coefficient of variation\

Rule 3: Estimates of differences between aggregates or percentages

The standard error of a difference between two estimates is approximately equal to the square root of the sum of squares of each standard error considered separately. That is, the standard error of a difference $\left(\hat{d} = \hat{X}_1 - \hat{X}_2\right)$ is:

$$\sigma_{\hat{d}} = \sqrt{\left(\hat{X}_1\alpha_1\right)^2 + \left(\hat{X}_2\alpha_2\right)^2}$$

where $\hat{X}_1$ is estimate 1, $\hat{X}_2$ is estimate 2, and α_1 and α_2 are the coefficients of variation of $\hat{X}_1$ and $\hat{X}_2$ respectively. The coefficient of variation of $\hat{d}$ is given by $\sigma_{\hat{d}} / \hat{d}$. This formula is accurate for the difference between separate and uncorrelated characteristics, but is only approximate otherwise.

Rule 4: Estimates of ratios

In the case where the numerator is a subset of the denominator, the ratio should be converted to a percentage and Rule 2 applied. This would apply, for example, to the case where the denominator is the number of females aged 16 to 65 and the numerator is the number of females aged 16 to 65 who have completed post-secondary education.

In the case where the numerator is not a subset of the denominator, as for example, the ratio of the number of females aged 16 to 65 who have completed post-secondary education as compared to the number of males aged 16 to 65 who have completed post-secondary education, the standard error of the ratio of the estimates is approximately equal to the square root of the sum of squares of each coefficient of variation considered separately multiplied by $\hat{R}$. That is, the standard error of a ratio $\left(\hat{R} = \hat{X}_1 / \hat{X}_2\right)$ is:

$$\sigma_{\hat{R}} = \hat{R}\sqrt{\alpha_1^{\,2} + \alpha_2^{\,2}}$$

where α_1 and α_2 are the coefficients of variation of $\hat{X}_1$ and $\hat{X}_2$ respectively. The coefficient of variation of $\hat{R}$ is given by $\sigma_{\hat{R}} / \hat{R}$. The formula will tend to overstate the error if $\hat{X}_1$ and $\hat{X}_2$ are positively correlated and understate the error if $\hat{X}_1$ and $\hat{X}_2$ are negatively correlated.

Rule 5: Estimates of differences of ratios

In this case, Rules 3 and 4 are combined. The CVs for the two ratios are first determined using Rule 4, and then the CV of their difference is found using Rule 3.

12.4 Examples of using the coefficient of variation tables for categorical estimates

The following examples based on the IALSS are included to assist users in applying the foregoing rules.

Example 1: Estimates of numbers of persons possessing a characteristic (Aggregates)

Suppose that a user estimates that 7,114,899 females aged between 16 and 65 have completed post-secondary education. How does the user determine the coefficient of variation of this estimate?

1) Refer to the coefficient of variation table for Canada in Section 12.9.

2) The estimated aggregate 7,114,899 does not appear in the left-hand column (the “Numerator of Percentage” column), so it is necessary to use the figure closest to it, namely 7,000,000.

3) The coefficient of variation for an estimated aggregate is found by referring to the first non-asterisk entry on that row, namely, 1.6%.

4) So the approximate coefficient of variation of the estimate is 1.6%. The finding that there were 7,114,899 females aged between 16 and 16 who have completed post-secondary education is publishable with no qualifications.

Example 2: Estimates of proportions or percentages of persons possessing a characteristic

Suppose that the user estimates that 7,114,899 / 11,555,360 = 61.6% of females aged between 16 and 65 have completed post-secondary education. How does the user determine the coefficient of variation of this estimate?

1) Refer to the coefficient of variation table for CANADA in Section 12.9.

2) Because the estimate is a percentage which is based on a subset of the total population (i.e., females who reported their education level), it is necessary to use both the percentage (61.6%) and the numerator portion of the percentage (7,114,899) in determining the coefficient of variation.

3) The numerator, 7,114,899, does not appear in the left-hand column (the “Numerator of Percentage” column) so it is necessary to use the figure closest to it, namely 7,000,000. Similarly, the percentage estimate does not appear as any of the column headings, so it is necessary to use the percentage closest to it, 70.0%.

4) The figure at the intersection of the row and column used, namely 1.1% is the coefficient of variation to be used.

5) So the approximate coefficient of variation of the estimate is 1.1%. The finding that 61.6% of females aged 16 to 65 have completed post-secondary education can be published with no qualifications.

Example 3: Estimates of differences between aggregates or percentages

Suppose that a user estimates that 7,114,899 / 11,555,360 = 61.6% of females aged between 16 and 65 who have completed post-secondary education, while 6,848,966 / 11,546,960 = 59.3% of males aged between 16 and 65 who have completed post-secondary education. How does the user determine the coefficient of variation of the difference between these two estimates?

1) Using the CANADA coefficient of variation table in the same manner as described in Example 2 gives the CV of the estimate for women as 1.1%, and the CV of the estimate for men as 1.4%.

2) Using Rule 3, the standard error of a difference $\left(\hat{d} = \hat{X}_1 - \hat{X}_2\right)$ is:

$$\sigma_{\hat{d}} = \sqrt{\left(\hat{X}_1\alpha_1\right)^2 + \left(\hat{X}_2\alpha_2\right)^2}$$

where $\hat{X}_1$ is estimate 1 (women), $\hat{X}_2$ is estimate 2 (men), and α_1 and α_2 are the coefficients of variation of $\hat{X}_1$ and $\hat{X}_2$ respectively.

That is, the standard error of the difference $\hat{d} =$ 0.593 – 0.616 = -0.023 is:

$$\begin{aligned}\sigma_{\hat{d}} &= \sqrt{[(0.616)(0.011)]^2 + [(0.593)(0.014)]^2}\\ &= \sqrt{(0.00006896) + (0.00004587)}\\ &= 0.0107\end{aligned}$$

3) The coefficient of variation of $\hat{d}$ is given by $\sigma_{\hat{d}} / \hat{d} =$ 0.0107 / 0.023 = 0.475

4) So the approximate coefficient of variation of the difference between the estimates is 47.5%. The difference between the estimates is considered unacceptable and Statistics Canada recommends this estimate not be released. However, should the user choose to do so, the estimate should be flagged with the letter U (or some similar identifier) and be accompanied by a warning to caution subsequent users about the high levels of error associated with the estimate.

Example 4: Estimates of ratios

Suppose that the user estimates that 7,196,746 persons aged between 16 and 65 in Ontario read articles in newspapers, magazine or newsletters at least once a week or everyday (H_Q01c=4 or 5), while 4,315,876 persons aged between 16 and 65 in Ontario read or use information from books, fiction or non-fiction at least once a week or everyday (H_Q01e=4 or 5). The user is interested in comparing the estimate of persons who read newspapers versus persons who read books in the form of a ratio. How does the user determine the coefficient of variation of this estimate?

1) First of all, this estimate is a ratio estimate, where the numerator of the estimate ($\hat{X}_1$) is the number of persons aged between 16 and 65 in Ontario who read or use information from newspapers at least once a week or everyday. The denominator of the estimate ($\hat{X}_2$) is the number of persons in Ontario who read or use information from books at least once a week or everyday.

2) Refer to the coefficient of variation table for Ontario in Section 12.9.

3) The numerator of this ratio estimate is 7,196,746. The figure closest to it is 7,000,000. The coefficient of variation for this estimate is found by referring to the first non-asterisk entry on that row, namely, 0.7%.

4) The denominator of this ratio estimate is 4,315,876. The figure closest to it is 4,000,000. The coefficient of variation for this estimate is found by referring to the first non-asterisk entry on that row, namely, 2.2%

5) So the approximate coefficient of variation of the ratio estimate is given by Rule 4, which is:

$$\alpha_{\hat{R}} = \sqrt{\alpha_1^{\ 2} + \alpha_2^{\ 2}}$$

where α_1 and α_2 are the coefficients of variation of $\hat{X}_1$ and $\hat{X}_2$ respectively.

That is:

$$\begin{aligned}\alpha_{\hat{R}} &= \sqrt{(0.007)^2 + (0.022)^2} \\ &= \sqrt{0.000049 + 0.000484} \\ &= 0.023\end{aligned}$$

6) The obtained ratio of persons in Ontario who read or use information from newspapers versus persons in Ontario who read or use information from books is 7,196,746 / 4,315,876 which is 1.668. The coefficient of variation of this estimate is 2.3%, which makes the estimate releasable with no qualifications.

Example 5: Estimates of differences of ratios

Suppose that the user estimates that the ratio of males in Ontario who read articles in newspapers at least once a week or everyday to males in Ontario who read or use information from books at least once a week or everyday is 2.24, while it is 1.33 for females. The user is interested in comparing the two ratios to see if there is a statistical difference between them. How does the user determine the coefficient of variation of the difference?

1) First calculate the approximate coefficient of variation for the male ratio ($\hat{R}_1$) and the female ratio ($\hat{R}_2$) as in Example 4. The approximate CV for the male ratio is 5.0% and 3.6% for females.

2) Using Rule 3, the standard error of a difference ($\hat{d} = \hat{R}_1 - \hat{R}_2$) is:

$$\sigma_{\hat{d}} = \sqrt{\left(\hat{R}_1\alpha_1\right)^2 + \left(\hat{R}_2\alpha_2\right)^2}$$

where α_1 and α_2 are the coefficients of variation of $\hat{R}_1$ and $\hat{R}_2$ respectively. That is, the standard error of the difference $\hat{d}$ = 2.24 – 1.33 = 0.91 is:

$$\begin{aligned}\sigma_{\hat{d}} &= \sqrt{[(2.24)(0.05)]^2 + [(1.33)(0.036)]^2} \\ &= \sqrt{(0.0112) + (0.00474)} \\ &= 0.122\end{aligned}$$

3) The coefficient of variation of $\hat{d}$ is given by $\sigma_{\hat{d}} / \hat{d}$ = 0.122 / (0.91) = 0.134.

4) So the approximate coefficient of variation of the difference between the estimates is 13.4% which makes the estimate releasable with no qualifications.

12.5 How to use the coefficient of variation tables to obtain confidence limits

Although coefficients of variation are widely used, a more intuitively meaningful measure of sampling error is the confidence interval of an estimate. A confidence interval constitutes a statement on the level of confidence that the true value for the population lies within a specified range of values. For example a 95% confidence interval can be described as follows:

If sampling of the population is repeated indefinitely, each sample leading to a new confidence interval for an estimate, then in 95% of the samples the interval will cover the true population value.

Using the standard error of an estimate, confidence intervals for estimates may be obtained under the assumption that under repeated sampling of the population, the various estimates obtained for a population characteristic are normally distributed about the true population value. Under this assumption, the chances are about 68 out of 100 that the difference between a sample estimate and the true population value would be less than one standard error, about 95 out of 100 that the difference would be less than two standard errors, and about 99 out of 100 that the difference would be less than three standard errors. These different degrees of confidence are referred to as the confidence levels.

Confidence intervals for an estimate, $\hat{X}$, are generally expressed as two numbers, one below the estimate and one above the estimate, as $(\hat{X}-k, \hat{X}+k)$ where k is determined depending upon the level of confidence desired and the sampling error of the estimate.

Confidence intervals for an estimate can be calculated directly from the Approximate Sampling Variability Tables by first determining from the appropriate table the coefficient of variation of the estimate $\hat{X}$, and then using the following formula to convert to a confidence interval ($CI_{\hat{x}}$):

$$CI_{\hat{x}} = \left(\hat{X} - t\hat{X}\alpha_{\hat{x}},\ \hat{X} + t\hat{X}\alpha_{\hat{x}}\right)$$

where $\alpha_{\hat{x}}$ is the determined coefficient of variation of $\hat{X}$, and

$t = 1$ if a 68% confidence interval is desired;

$t = 1.6$ if a 90% confidence interval is desired;

$t = 2$ if a 95% confidence interval is desired;

$t = 2.6$ if a 99% confidence interval is desired.

Note: Release guidelines which apply to the estimate also apply to the confidence interval. For example, if the estimate is not releasable, then the confidence interval is not releasable either.

Example of using the coefficient of variation tables to obtain confidence limits

A 95% confidence interval for the estimated proportion of females who have completed post-secondary education (from Example 2, Section 12.4) would be calculated as follows:

$\hat{X}$ = 61.6 % (or expressed as a proportion 0.616)

t = 2

$\alpha_{\hat{x}}$ = 1.1 % (0.011 expressed as a proportion) is the coefficient of variation of this estimate as determined from the tables.

$CI_{\hat{x}}$ = {0. 616– (2) (0.616) (0.11). 0.616 + (2) (0.616) (0.011)}

$CI_{\hat{x}}$ = {0. 616– 0.0135. 0. 616+ 0. 0135}

$CI_{\hat{x}}$ = {0.602. 0.629}

With 95% confidence it can be said that between 60.2% and 62.9% of females have completed post-secondary education.

12.6 How to use the coefficient of variation tables to do a T-test

Standard errors may also be used to perform hypothesis testing, a procedure for distinguishing between population parameters using sample estimates. The sample estimates can be numbers, averages, percentages, ratios, etc. Tests may be performed at various levels of significance, where a level of significance is the probability of concluding that the characteristics are different when, in fact, they are identical.

Let $\hat{X}_1$ and $\hat{X}_2$ be sample estimates for two characteristics of interest. Let the standard error on the difference $\hat{X}_1 - \hat{X}_2$ be $\sigma_{\hat{d}}$.

If $t = \frac{\hat{X}_1 - \hat{X}_2}{\sigma_{\hat{d}}}$ is between -2 and 2, then no conclusion about the difference between the characteristics is justified at the 5% level of significance. If however, this ratio is smaller than -2 or larger than +2, the observed difference is significant at the 0.05 level. That is to say that the difference between the estimates is significant.

Example of using the coefficient of variation tables to do a T-test

Let us suppose that the user wishes to test, at 5% level of significance, the hypothesis that there is no difference between the proportion of females who have completed post-secondary education and the proportion of males who have completed post-secondary education. From Example 3, Section 12.4, the standard error of the difference between these two estimates was found to be 0.023. Hence,

$$t = \frac{\hat{X}_1 - \hat{X}_2}{\sigma_{\hat{d}}} = \frac{0.616 - 0.593}{0.0107} = \frac{0.023}{0.0107} = 2.11$$

Since t = 2.11 is not between -2 and 2, we can consider that the difference is significant at a 5% threshold. The percentage of female that have obtained a post-secondary diploma is significantly different than the males.

12.7 Coefficients of variation for quantitative estimates

For quantitative estimates, special tables would have to be produced to determine their sampling error. Since most of the variables for the PIAAC are primarily categorical in nature, this has not been done.

As a general rule, however, the coefficient of variation of a quantitative total will be larger than the coefficient of variation of the corresponding category estimate (i.e., the estimate of the number of persons contributing to the quantitative estimate). If the corresponding category estimate is not releasable, the quantitative estimate will not be either. For example, the coefficient of variation of the total number of hours spent in a program of studies would be greater than the coefficient of variation of the corresponding proportion of persons enrolled in a program of studies. Hence, if the coefficient of variation of the proportion is unacceptable (making the proportion not releasable), then the coefficient of variation of the corresponding quantitative estimate will also be unacceptable (making the quantitative estimate not releasable). Coefficients of variation of such estimates can be derived using the 80 jackknife replicate weights, as described in Section 11.2.

12.8 Coefficients of Variation for Skill Level Estimates

As explained in section 11.2.3, the ten plausible values used for estimating skill levels add another component to the variance, called the imputation variance. A factor was calculated which should be used to inflate the CVs obtained from the CV tables for estimates of skill level. The inflation factor was determined by first calculating the percentage contribution of the imputation variance to the overall variance for a large number of skill level estimates and then choosing the 75th percentile.

Rule 6: Estimates of Skill Level

The CVs of skill level estimates are approximated by multiplying the value obtained from the CV table by a factor of 1.4. This rule is to be used in conjunction with Rules 1 to 5 in section 12.3. For estimates involving more than one variable, Rule 6 is used first to approximate the CV of each skill level variable, and then Rules 1 to 5 are applied to combine the CVs and calculate an overall CV. Two examples are provided below.

Example 6.1: Estimates of Proportions with Skill Levels (Rule 2 & Rule 6)

Suppose that a user estimates that 1,887,142 / 11,590,762 = 16.3% of males aged between 16 and 65 obtained a score in Level 1 (i.e. less than 226) in literacy. How does the user determine the coefficient of variation of this proportion?

Refer to the coefficient of variation table for CANADA in section 0.

Using Rule 2, the CV obtained from the CV table is 3.4%.

Multiplying 3.4% by the inflation factor of 1.4 gives a CV of 4.8%.

So the approximate coefficient of variation of the proportion is 4.8%, which makes the estimate releasable with no qualifications.

Example 6.2: Estimates of Ratios with Skill Levels (Rule 6 & Rule 4)

Suppose that a user estimates that 3,830,422 persons obtained a score in Level 1 (i.e., less than 226) in literacy, while 3,548,580 persons aged between 16 and 65 completed less than high school or had no formal education. The user is interested in comparing the estimate of these two quantities in the form of a ratio. How does the user determine the coefficient of variation of this ratio?

1) First of all, this estimate is a ratio estimate, where the numerator of the estimate ($\hat{X}_1$) is the number of persons who obtained a score in the Level 1 prose literacy category. The denominator of the estimate ($\hat{X}_2$) is the number of persons who completed less than high school.

2) Refer to the coefficient of variation table for CANADA in section 0.

3) The numerator of this ratio estimate is 3,830,422. The CV obtained from the CV tables for the

numerator is 2.3%. Since the numerator is a skill level, Rule 6 applies; the CV is multiplied by a factor of 1.4, which yields an approximate CV for the numerator of 3.2%.

4) The denominator of this ratio estimate is 3,548,580. The CV obtained from the CV tables for the denominator is 2.3%. The denominator is not a skill level so Rule 6 does not apply.

5) The approximate coefficient of variation of the ratio estimate is given by Rule 4, which is:

$$\alpha_{\hat{R}} = \sqrt{\alpha_1^{\ 2} + \alpha_2^{\ 2}}$$

where α_1 and α_2 are the coefficients of variation of $\hat{X}_1$ and $\hat{X}_2$ respectively. That is:

$$\alpha_{\hat{R}} = \sqrt{(0.032)^2 + (0.023)^2} = 0.0396$$

6) So the ratio 3,830,422 / 3,548,580 = 1.08 has an approximate coefficient of variation of 4.0%, which makes the estimate releasable with no qualifications.

12.9 Coefficient of variation tables

Programme for the International Assessment of Adult Competencies, 2012

Approximate Sampling Variability Tables - Canada

NUMERATOR OF PERCENTAGE ('000)	ESTIMATED PERCENTAGE 0.1%	1.0%	2.0%	5.0%	10.0%	15.0%	20.0%	25.0%	30.0%	35.0%	40.0%	50.0%	70.0%	90.0%
1	163.6	162.8	162.0	159.5	155.2	150.9	146.4	141.7	136.9	131.9	126.8	115.7	89.6	51.7
2	115.7	115.1	114.5	112.8	109.8	106.7	103.5	100.2	96.8	93.3	89.6	81.8	63.4	36.6
3	94.4	94.0	93.5	92.1	89.6	87.1	84.5	81.8	79.0	76.2	73.2	66.8	51.7	29.9
4	81.8	81.4	81.0	79.7	77.6	75.4	73.2	70.9	68.5	66.0	63.4	57.9	44.8	25.9
5	73.1	72.8	72.4	71.3	69.4	67.5	65.5	63.4	61.2	59.0	56.7	51.7	40.1	23.1
6	66.8	66.5	66.1	65.1	63.4	61.6	59.8	57.9	55.9	53.9	51.7	47.2	36.6	21.1
7	61.8	61.5	61.2	60.3	58.7	57.0	55.3	53.6	51.7	49.9	47.9	43.7	33.9	19.6
8	57.8	57.6	57.3	56.4	54.9	53.3	51.7	50.1	48.4	46.6	44.8	40.9	31.7	18.3
9	54.5	54.3	54.0	53.2	51.7	50.3	48.8	47.2	45.6	44.0	42.3	38.6	29.9	17.2
10	51.7	51.5	51.2	50.4	49.1	47.7	46.3	44.8	43.3	41.7	40.1	36.6	28.3	16.4
11	49.3	49.1	48.8	48.1	46.8	45.5	44.1	42.7	41.3	39.8	38.2	34.9	27.0	15.6
12	47.2	47.0	46.8	46.0	44.8	43.6	42.3	40.9	39.5	38.1	36.6	33.4	25.9	14.9
13	45.4	45.2	44.9	44.2	43.1	41.8	40.6	39.3	38.0	36.6	35.2	32.1	24.9	14.4
14	43.7	43.5	43.3	42.6	41.5	40.3	39.1	37.9	36.6	35.3	33.9	30.9	24.0	13.8
15	42.2	42.0	41.8	41.2	40.1	39.0	37.8	36.6	35.4	34.1	32.7	29.9	23.1	13.4
16	40.9	40.7	40.5	39.9	38.8	37.7	36.6	35.4	34.2	33.0	31.7	28.9	22.4	12.9
17	39.7	39.5	39.3	38.7	37.7	36.6	35.5	34.4	33.2	32.0	30.7	28.1	21.7	12.6
18	38.6	38.4	38.2	37.6	36.6	35.6	34.5	33.4	32.3	31.1	29.9	27.3	21.1	12.2
19	37.5	37.4	37.2	36.6	35.6	34.6	33.6	32.5	31.4	30.3	29.1	26.5	20.6	11.9
20	36.6	36.4	36.2	35.7	34.7	33.7	32.7	31.7	30.6	29.5	28.3	25.9	20.0	11.6
21	35.7	35.5	35.4	34.8	33.9	32.9	31.9	30.9	29.9	28.8	27.7	25.3	19.6	11.3
22	34.9	34.7	34.5	34.0	33.1	32.2	31.2	30.2	29.2	28.1	27.0	24.7	19.1	11.0
23	34.1	34.0	33.8	33.3	32.4	31.5	30.5	29.6	28.5	27.5	26.4	24.1	18.7	10.8
24	********	33.2	33.1	32.6	31.7	30.8	29.9	28.9	27.9	26.9	25.9	23.6	18.3	10.6
25	********	32.6	32.4	31.9	31.0	30.2	29.3	28.3	27.4	26.4	25.4	23.1	17.9	10.3
30	********	29.7	29.6	29.1	28.3	27.5	26.7	25.9	25.0	24.1	23.1	21.1	16.4	9.4
35	********	27.5	27.4	27.0	26.2	25.5	24.7	24.0	23.1	22.3	21.4	19.6	15.2	8.7
40	********	25.7	25.6	25.2	24.5	23.9	23.1	22.4	21.6	20.9	20.0	18.3	14.2	8.2
45	********	24.3	24.1	23.8	23.1	22.5	21.8	21.1	20.4	19.7	18.9	17.2	13.4	7.7
50	********	23.0	22.9	22.6	22.0	21.3	20.7	20.0	19.4	18.7	17.9	16.4	12.7	7.3
55	********	22.0	21.8	21.5	20.9	20.3	19.7	19.1	18.5	17.8	17.1	15.6	12.1	7.0
60	********	21.0	20.9	20.6	20.0	19.5	18.9	18.3	17.7	17.0	16.4	14.9	11.6	6.7
65	********	20.2	20.1	19.8	19.3	18.7	18.2	17.6	17.0	16.4	15.7	14.4	11.1	6.4
70	********	19.5	19.4	19.1	18.6	18.0	17.5	16.9	16.4	15.8	15.2	13.8	10.7	6.2
75	********	18.8	18.7	18.4	17.9	17.4	16.9	16.4	15.8	15.2	14.6	13.4	10.3	6.0
80	********	18.2	18.1	17.8	17.4	16.9	16.4	15.8	15.3	14.8	14.2	12.9	10.0	5.8
85	********	17.7	17.6	17.3	16.8	16.4	15.9	15.4	14.9	14.3	13.7	12.6	9.7	5.6
90	********	17.2	17.1	16.8	16.4	15.9	15.4	14.9	14.4	13.9	13.4	12.2	9.4	5.5
95	********	16.7	16.6	16.4	15.9	15.5	15.0	14.5	14.0	13.5	13.0	11.9	9.2	5.3
100	********	16.3	16.2	15.9	15.5	15.1	14.6	14.2	13.7	13.2	12.7	11.6	9.0	5.2
125	********	14.6	14.5	14.3	13.9	13.5	13.1	12.7	12.2	11.8	11.3	10.3	8.0	4.6
150	********	13.3	13.2	13.0	12.7	12.3	12.0	11.6	11.2	10.8	10.3	9.4	7.3	4.2
200	********	11.5	11.5	11.3	11.0	10.7	10.3	10.0	9.7	9.3	9.0	8.2	6.3	3.7
250	****************		10.2	10.1	9.8	9.5	9.3	9.0	8.7	8.3	8.0	7.3	5.7	3.3
300	****************		9.4	9.2	9.0	8.7	8.5	8.2	7.9	7.6	7.3	6.7	5.2	3.0
350	****************		8.7	8.5	8.3	8.1	7.8	7.6	7.3	7.1	6.8	6.2	4.8	2.8
400	****************		8.1	8.0	7.8	7.5	7.3	7.1	6.8	6.6	6.3	5.8	4.5	2.6
450	****************		7.6	7.5	7.3	7.1	6.9	6.7	6.5	6.2	6.0	5.5	4.2	2.4
500	************************			7.1	6.9	6.7	6.5	6.3	6.1	5.9	5.7	5.2	4.0	2.3
750	************************			5.8	5.7	5.5	5.3	5.2	5.0	4.8	4.6	4.2	3.3	1.9
1,000	************************			5.0	4.9	4.8	4.6	4.5	4.3	4.2	4.0	3.7	2.8	1.6
1,500	********************************				4.0	3.9	3.8	3.7	3.5	3.4	3.3	3.0	2.3	1.3
2,000	********************************				3.5	3.4	3.3	3.2	3.1	3.0	2.8	2.6	2.0	1.2
3,000	**					2.8	2.7	2.6	2.5	2.4	2.3	2.1	1.6	0.9
4,000	**						2.3	2.2	2.2	2.1	2.0	1.8	1.4	0.8
5,000	**							2.0	1.9	1.9	1.8	1.6	1.3	0.7
6,000	**								1.8	1.7	1.6	1.5	1.2	0.7
7,000	**									1.6	1.5	1.4	1.1	0.6
8,000	**									1.5	1.4	1.3	1.0	0.6
9,000	**										1.3	1.2	0.9	0.5
10,000	**											1.2	0.9	0.5
12,500	**												0.8	0.5
15,000	**												0.7	0.4
20,000	**													0.4

NOTE: For correct usage of these tables, please refer to the microdata documentation.

Programme for the International Assessment of Adult Competencies, 2012

Approximate Sampling Variability Tables - Atlantic Provinces

NUMERATOR OF PERCENTAGE ('000)	ESTIMATED PERCENTAGE 0.1%	1.0%	2.0%	5.0%	10.0%	15.0%	20.0%	25.0%	30.0%	35.0%	40.0%	50.0%	70.0%	90.0%
1	68.0	67.7	67.4	66.3	64.6	62.7	60.9	58.9	56.9	54.9	52.7	48.1	37.3	21.5
2	********	47.9	47.6	46.9	45.6	44.4	43.0	41.7	40.3	38.8	37.3	34.0	26.4	15.2
3	********	39.1	38.9	38.3	37.3	36.2	35.1	34.0	32.9	31.7	30.4	27.8	21.5	12.4
4	********	33.9	33.7	33.2	32.3	31.4	30.4	29.5	28.5	27.4	26.4	24.1	18.6	10.8
5	********	30.3	30.1	29.7	28.9	28.1	27.2	26.4	25.5	24.5	23.6	21.5	16.7	9.6
6	********	27.6	27.5	27.1	26.4	25.6	24.8	24.1	23.2	22.4	21.5	19.6	15.2	8.8
7	********	25.6	25.5	25.1	24.4	23.7	23.0	22.3	21.5	20.7	19.9	18.2	14.1	8.1
8	********	23.9	23.8	23.4	22.8	22.2	21.5	20.8	20.1	19.4	18.6	17.0	13.2	7.6
9	********	22.6	22.5	22.1	21.5	20.9	20.3	19.6	19.0	18.3	17.6	16.0	12.4	7.2
10	********	21.4	21.3	21.0	20.4	19.8	19.2	18.6	18.0	17.3	16.7	15.2	11.8	6.8
11	********	20.4	20.3	20.0	19.5	18.9	18.4	17.8	17.2	16.5	15.9	14.5	11.2	6.5
12	********	19.5	19.4	19.1	18.6	18.1	17.6	17.0	16.4	15.8	15.2	13.9	10.8	6.2
13	********	18.8	18.7	18.4	17.9	17.4	16.9	16.3	15.8	15.2	14.6	13.3	10.3	6.0
14	********	18.1	18.0	17.7	17.3	16.8	16.3	15.7	15.2	14.7	14.1	12.9	10.0	5.8
15	********	17.5	17.4	17.1	16.7	16.2	15.7	15.2	14.7	14.2	13.6	12.4	9.6	5.6
16	********	********	16.8	16.6	16.1	15.7	15.2	14.7	14.2	13.7	13.2	12.0	9.3	5.4
17	********	********	16.3	16.1	15.7	15.2	14.8	14.3	13.8	13.3	12.8	11.7	9.0	5.2
18	********	********	15.9	15.6	15.2	14.8	14.3	13.9	13.4	12.9	12.4	11.3	8.8	5.1
19	********	********	15.5	15.2	14.8	14.4	14.0	13.5	13.1	12.6	12.1	11.0	8.6	4.9
20	********	********	15.1	14.8	14.4	14.0	13.6	13.2	12.7	12.3	11.8	10.8	8.3	4.8
21	********	********	14.7	14.5	14.1	13.7	13.3	12.9	12.4	12.0	11.5	10.5	8.1	4.7
22	********	********	14.4	14.1	13.8	13.4	13.0	12.6	12.1	11.7	11.2	10.3	7.9	4.6
23	********	********	14.0	13.8	13.5	13.1	12.7	12.3	11.9	11.4	11.0	10.0	7.8	4.5
24	********	********	13.8	13.5	13.2	12.8	12.4	12.0	11.6	11.2	10.8	9.8	7.6	4.4
25	********	********	13.5	13.3	12.9	12.5	12.2	11.8	11.4	11.0	10.5	9.6	7.5	4.3
30	********	********	12.3	12.1	11.8	11.5	11.1	10.8	10.4	10.0	9.6	8.8	6.8	3.9
35	********	********	********	11.2	10.9	10.6	10.3	10.0	9.6	9.3	8.9	8.1	6.3	3.6
40	********	********	********	10.5	10.2	9.9	9.6	9.3	9.0	8.7	8.3	7.6	5.9	3.4
45	********	********	********	9.9	9.6	9.4	9.1	8.8	8.5	8.2	7.9	7.2	5.6	3.2
50	********	********	********	9.4	9.1	8.9	8.6	8.3	8.1	7.8	7.5	6.8	5.3	3.0
55	********	********	********	8.9	8.7	8.5	8.2	7.9	7.7	7.4	7.1	6.5	5.0	2.9
60	********	********	********	8.6	8.3	8.1	7.9	7.6	7.3	7.1	6.8	6.2	4.8	2.8
65	********	********	********	8.2	8.0	7.8	7.5	7.3	7.1	6.8	6.5	6.0	4.6	2.7
70	********	********	********	7.9	7.7	7.5	7.3	7.0	6.8	6.6	6.3	5.8	4.5	2.6
75	********	********	********	7.7	7.5	7.2	7.0	6.8	6.6	6.3	6.1	5.6	4.3	2.5
80	********	********	********	********	7.2	7.0	6.8	6.6	6.4	6.1	5.9	5.4	4.2	2.4
85	********	********	********	********	7.0	6.8	6.6	6.4	6.2	6.0	5.7	5.2	4.0	2.3
90	********	********	********	********	6.8	6.6	6.4	6.2	6.0	5.8	5.6	5.1	3.9	2.3
95	********	********	********	********	6.6	6.4	6.2	6.0	5.8	5.6	5.4	4.9	3.8	2.2
100	********	********	********	********	6.5	6.3	6.1	5.9	5.7	5.5	5.3	4.8	3.7	2.2
125	********	********	********	********	5.8	5.6	5.4	5.3	5.1	4.9	4.7	4.3	3.3	1.9
150	********	********	********	********	5.3	5.1	5.0	4.8	4.6	4.5	4.3	3.9	3.0	1.8
200	********	********	********	********	********	4.4	4.3	4.2	4.0	3.9	3.7	3.4	2.6	1.5
250	********	********	********	********	********	********	3.8	3.7	3.6	3.5	3.3	3.0	2.4	1.4
300	********	********	********	********	********	********	3.5	3.4	3.3	3.2	3.0	2.8	2.2	1.2
350	********	********	********	********	********	********	********	3.1	3.0	2.9	2.8	2.6	2.0	1.2
400	********	********	********	********	********	********	********	********	2.8	2.7	2.6	2.4	1.9	1.1
450	********	********	********	********	********	********	********	********	2.7	2.6	2.5	2.3	1.8	1.0
500	********	********	********	********	********	********	********	********	********	2.5	2.4	2.2	1.7	1.0
750	********	********	********	********	********	********	********	********	********	********	********	1.8	1.4	0.8
1,000	********	********	********	********	********	********	********	********	********	********	********	********	1.2	0.7

NOTE: For correct usage of these tables, please refer to the microdata documentation.

Programme for the International Assessment of Adult Competencies, 2012

Approximate Sampling Variability Tables - Quebec

NUMERATOR OF PERCENTAGE ('000)	ESTIMATED PERCENTAGE													
	0.1%	1.0%	2.0%	5.0%	10.0%	15.0%	20.0%	25.0%	30.0%	35.0%	40.0%	50.0%	70.0%	90.0%
1	110.6	110.1	109.5	107.8	105.0	102.0	98.9	95.8	92.6	89.2	85.7	78.2	60.6	35.0
2	78.2	77.8	77.4	76.2	74.2	72.1	70.0	67.7	65.4	63.1	60.6	55.3	42.8	24.7
3	63.8	63.6	63.2	62.3	60.6	58.9	57.1	55.3	53.4	51.5	49.5	45.2	35.0	20.2
4	55.3	55.0	54.8	53.9	52.5	51.0	49.5	47.9	46.3	44.6	42.8	39.1	30.3	17.5
5	49.5	49.2	49.0	48.2	46.9	45.6	44.3	42.8	41.4	39.9	38.3	35.0	27.1	15.6
6	********	44.9	44.7	44.0	42.8	41.6	40.4	39.1	37.8	36.4	35.0	31.9	24.7	14.3
7	********	41.6	41.4	40.8	39.7	38.6	37.4	36.2	35.0	33.7	32.4	29.6	22.9	13.2
8	********	38.9	38.7	38.1	37.1	36.1	35.0	33.9	32.7	31.5	30.3	27.7	21.4	12.4
9	********	36.7	36.5	35.9	35.0	34.0	33.0	31.9	30.9	29.7	28.6	26.1	20.2	11.7
10	********	34.8	34.6	34.1	33.2	32.3	31.3	30.3	29.3	28.2	27.1	24.7	19.2	11.1
11	********	33.2	33.0	32.5	31.6	30.8	29.8	28.9	27.9	26.9	25.8	23.6	18.3	10.5
12	********	31.8	31.6	31.1	30.3	29.4	28.6	27.7	26.7	25.7	24.7	22.6	17.5	10.1
13	********	30.5	30.4	29.9	29.1	28.3	27.4	26.6	25.7	24.7	23.8	21.7	16.8	9.7
14	********	29.4	29.3	28.8	28.0	27.3	26.4	25.6	24.7	23.8	22.9	20.9	16.2	9.3
15	********	28.4	28.3	27.8	27.1	26.3	25.5	24.7	23.9	23.0	22.1	20.2	15.6	9.0
16	********	27.5	27.4	27.0	26.2	25.5	24.7	24.0	23.1	22.3	21.4	19.6	15.1	8.7
17	********	26.7	26.6	26.2	25.5	24.7	24.0	23.2	22.4	21.6	20.8	19.0	14.7	8.5
18	********	25.9	25.8	25.4	24.7	24.0	23.3	22.6	21.8	21.0	20.2	18.4	14.3	8.2
19	********	25.3	25.1	24.7	24.1	23.4	22.7	22.0	21.2	20.5	19.7	17.9	13.9	8.0
20	********	24.6	24.5	24.1	23.5	22.8	22.1	21.4	20.7	19.9	19.2	17.5	13.5	7.8
21	********	24.0	23.9	23.5	22.9	22.3	21.6	20.9	20.2	19.5	18.7	17.1	13.2	7.6
22	********	23.5	23.3	23.0	22.4	21.7	21.1	20.4	19.7	19.0	18.3	16.7	12.9	7.5
23	********	23.0	22.8	22.5	21.9	21.3	20.6	20.0	19.3	18.6	17.9	16.3	12.6	7.3
24	********	22.5	22.4	22.0	21.4	20.8	20.2	19.6	18.9	18.2	17.5	16.0	12.4	7.1
25	********	22.0	21.9	21.6	21.0	20.4	19.8	19.2	18.5	17.8	17.1	15.6	12.1	7.0
30	********	20.1	20.0	19.7	19.2	18.6	18.1	17.5	16.9	16.3	15.6	14.3	11.1	6.4
35	********	18.6	18.5	18.2	17.7	17.2	16.7	16.2	15.6	15.1	14.5	13.2	10.2	5.9
40	********	17.4	17.3	17.0	16.6	16.1	15.6	15.1	14.6	14.1	13.5	12.4	9.6	5.5
45	********	16.4	16.3	16.1	15.6	15.2	14.8	14.3	13.8	13.3	12.8	11.7	9.0	5.2
50	********	15.6	15.5	15.2	14.8	14.4	14.0	13.5	13.1	12.6	12.1	11.1	8.6	4.9
55	****************		14.8	14.5	14.2	13.8	13.3	12.9	12.5	12.0	11.6	10.5	8.2	4.7
60	****************		14.1	13.9	13.5	13.2	12.8	12.4	11.9	11.5	11.1	10.1	7.8	4.5
65	****************		13.6	13.4	13.0	12.7	12.3	11.9	11.5	11.1	10.6	9.7	7.5	4.3
70	****************		13.1	12.9	12.5	12.2	11.8	11.5	11.1	10.7	10.2	9.3	7.2	4.2
75	****************		12.6	12.5	12.1	11.8	11.4	11.1	10.7	10.3	9.9	9.0	7.0	4.0
80	****************		12.2	12.1	11.7	11.4	11.1	10.7	10.3	10.0	9.6	8.7	6.8	3.9
85	****************		11.9	11.7	11.4	11.1	10.7	10.4	10.0	9.7	9.3	8.5	6.6	3.8
90	****************		11.5	11.4	11.1	10.8	10.4	10.1	9.8	9.4	9.0	8.2	6.4	3.7
95	****************		11.2	11.1	10.8	10.5	10.2	9.8	9.5	9.2	8.8	8.0	6.2	3.6
100	****************		11.0	10.8	10.5	10.2	9.9	9.6	9.3	8.9	8.6	7.8	6.1	3.5
125	************************			9.6	9.4	9.1	8.9	8.6	8.3	8.0	7.7	7.0	5.4	3.1
150	************************			8.8	8.6	8.3	8.1	7.8	7.6	7.3	7.0	6.4	4.9	2.9
200	************************			7.6	7.4	7.2	7.0	6.8	6.5	6.3	6.1	5.5	4.3	2.5
250	************************			6.8	6.6	6.5	6.3	6.1	5.9	5.6	5.4	4.9	3.8	2.2
300	********************************				6.1	5.9	5.7	5.5	5.3	5.1	4.9	4.5	3.5	2.0
350	********************************				5.6	5.5	5.3	5.1	4.9	4.8	4.6	4.2	3.2	1.9
400	********************************				5.2	5.1	4.9	4.8	4.6	4.5	4.3	3.9	3.0	1.7
450	********************************				4.9	4.8	4.7	4.5	4.4	4.2	4.0	3.7	2.9	1.6
500	********************************				4.7	4.6	4.4	4.3	4.1	4.0	3.8	3.5	2.7	1.6
750	**					3.7	3.6	3.5	3.4	3.3	3.1	2.9	2.2	1.3
1,000	**						3.1	3.0	2.9	2.8	2.7	2.5	1.9	1.1
1,500	**								2.4	2.3	2.2	2.0	1.6	0.9
2,000	**										1.9	1.7	1.4	0.8
3,000	**												1.1	0.6
4,000	**													0.6

NOTE: For correct usage of these tables, please refer to the microdata documentation.

Programme for the International Assessment of Adult Competencies, 2012

Approximate Sampling Variability Tables - Ontario

NUMERATOR OF PERCENTAGE ('000)	ESTIMATED PERCENTAGE 0.1%	1.0%	2.0%	5.0%	10.0%	15.0%	20.0%	25.0%	30.0%	35.0%	40.0%	50.0%	70.0%	90.0%
1	193.1	192.3	191.3	188.3	183.3	178.2	172.8	167.3	161.7	155.8	149.7	136.6	105.8	61.1
2	136.6	136.0	135.3	133.2	129.6	126.0	122.2	118.3	114.3	110.2	105.8	96.6	74.8	43.2
3	111.5	111.0	110.4	108.7	105.8	102.9	99.8	96.6	93.3	89.9	86.4	78.9	61.1	35.3
4	96.6	96.1	95.6	94.2	91.7	89.1	86.4	83.7	80.8	77.9	74.8	68.3	52.9	30.6
5	86.4	86.0	85.6	84.2	82.0	79.7	77.3	74.8	72.3	69.7	66.9	61.1	47.3	27.3
6	78.8	78.5	78.1	76.9	74.8	72.7	70.6	68.3	66.0	63.6	61.1	55.8	43.2	24.9
7	73.0	72.7	72.3	71.2	69.3	67.3	65.3	63.3	61.1	58.9	56.6	51.6	40.0	23.1
8	68.3	68.0	67.6	66.6	64.8	63.0	61.1	59.2	57.2	55.1	52.9	48.3	37.4	21.6
9	64.4	64.1	63.8	62.8	61.1	59.4	57.6	55.8	53.9	51.9	49.9	45.5	35.3	20.4
10	********	60.8	60.5	59.6	58.0	56.3	54.7	52.9	51.1	49.3	47.3	43.2	33.5	19.3
11	********	58.0	57.7	56.8	55.3	53.7	52.1	50.5	48.7	47.0	45.1	41.2	31.9	18.4
12	********	55.5	55.2	54.4	52.9	51.4	49.9	48.3	46.7	45.0	43.2	39.4	30.6	17.6
13	********	53.3	53.1	52.2	50.8	49.4	47.9	46.4	44.8	43.2	41.5	37.9	29.4	16.9
14	********	51.4	51.1	50.3	49.0	47.6	46.2	44.7	43.2	41.6	40.0	36.5	28.3	16.3
15	********	49.6	49.4	48.6	47.3	46.0	44.6	43.2	41.7	40.2	38.6	35.3	27.3	15.8
16	********	48.1	47.8	47.1	45.8	44.5	43.2	41.8	40.4	38.9	37.4	34.2	26.5	15.3
17	********	46.6	46.4	45.7	44.5	43.2	41.9	40.6	39.2	37.8	36.3	33.1	25.7	14.8
18	********	45.3	45.1	44.4	43.2	42.0	40.7	39.4	38.1	36.7	35.3	32.2	24.9	14.4
19	********	44.1	43.9	43.2	42.1	40.9	39.7	38.4	37.1	35.7	34.3	31.3	24.3	14.0
20	********	43.0	42.8	42.1	41.0	39.8	38.6	37.4	36.2	34.8	33.5	30.6	23.7	13.7
21	********	42.0	41.7	41.1	40.0	38.9	37.7	36.5	35.3	34.0	32.7	29.8	23.1	13.3
22	********	41.0	40.8	40.2	39.1	38.0	36.8	35.7	34.5	33.2	31.9	29.1	22.6	13.0
23	********	40.1	39.9	39.3	38.2	37.1	36.0	34.9	33.7	32.5	31.2	28.5	22.1	12.7
24	********	39.2	39.0	38.4	37.4	36.4	35.3	34.2	33.0	31.8	30.6	27.9	21.6	12.5
25	********	38.5	38.3	37.7	36.7	35.6	34.6	33.5	32.3	31.2	29.9	27.3	21.2	12.2
30	********	35.1	34.9	34.4	33.5	32.5	31.6	30.6	29.5	28.4	27.3	24.9	19.3	11.2
35	********	32.5	32.3	31.8	31.0	30.1	29.2	28.3	27.3	26.3	25.3	23.1	17.9	10.3
40	********	30.4	30.2	29.8	29.0	28.2	27.3	26.5	25.6	24.6	23.7	21.6	16.7	9.7
45	********	28.7	28.5	28.1	27.3	26.6	25.8	24.9	24.1	23.2	22.3	20.4	15.8	9.1
50	********	27.2	27.1	26.6	25.9	25.2	24.4	23.7	22.9	22.0	21.2	19.3	15.0	8.6
55	********	25.9	25.8	25.4	24.7	24.0	23.3	22.6	21.8	21.0	20.2	18.4	14.3	8.2
60	********	24.8	24.7	24.3	23.7	23.0	22.3	21.6	20.9	20.1	19.3	17.6	13.7	7.9
65	********	23.8	23.7	23.4	22.7	22.1	21.4	20.8	20.1	19.3	18.6	16.9	13.1	7.6
70	********	23.0	22.9	22.5	21.9	21.3	20.7	20.0	19.3	18.6	17.9	16.3	12.7	7.3
75	********	22.2	22.1	21.7	21.2	20.6	20.0	19.3	18.7	18.0	17.3	15.8	12.2	7.1
80	********	21.5	21.4	21.1	20.5	19.9	19.3	18.7	18.1	17.4	16.7	15.3	11.8	6.8
85	********	20.9	20.7	20.4	19.9	19.3	18.7	18.2	17.5	16.9	16.2	14.8	11.5	6.6
90	********	20.3	20.2	19.9	19.3	18.8	18.2	17.6	17.0	16.4	15.8	14.4	11.2	6.4
95	****************		19.6	19.3	18.8	18.3	17.7	17.2	16.6	16.0	15.4	14.0	10.9	6.3
100	****************		19.1	18.8	18.3	17.8	17.3	16.7	16.2	15.6	15.0	13.7	10.6	6.1
125	****************		17.1	16.8	16.4	15.9	15.5	15.0	14.5	13.9	13.4	12.2	9.5	5.5
150	****************		15.6	15.4	15.0	14.5	14.1	13.7	13.2	12.7	12.2	11.2	8.6	5.0
200	************************			13.3	13.0	12.6	12.2	11.8	11.4	11.0	10.6	9.7	7.5	4.3
250	************************			11.9	11.6	11.3	10.9	10.6	10.2	9.9	9.5	8.6	6.7	3.9
300	************************			10.9	10.6	10.3	10.0	9.7	9.3	9.0	8.6	7.9	6.1	3.5
350	************************			10.1	9.8	9.5	9.2	8.9	8.6	8.3	8.0	7.3	5.7	3.3
400	************************			9.4	9.2	8.9	8.6	8.4	8.1	7.8	7.5	6.8	5.3	3.1
450	************************			8.9	8.6	8.4	8.1	7.9	7.6	7.3	7.1	6.4	5.0	2.9
500	********************************				8.2	8.0	7.7	7.5	7.2	7.0	6.7	6.1	4.7	2.7
750	********************************				6.7	6.5	6.3	6.1	5.9	5.7	5.5	5.0	3.9	2.2
1,000	**					5.6	5.5	5.3	5.1	4.9	4.7	4.3	3.3	1.9
1,500	**						4.5	4.3	4.2	4.0	3.9	3.5	2.7	1.6
2,000	**							3.7	3.6	3.5	3.3	3.1	2.4	1.4
3,000	**									2.8	2.7	2.5	1.9	1.1
4,000	**											2.2	1.7	1.0
5,000	**												1.5	0.9
6,000	**												1.4	0.8
7,000	**													0.7
8,000	**													0.7

NOTE: For correct usage of these tables, please refer to the microdata documentation.

Programme for the International Assessment of Adult Competencies, 2012

NUMERATOR OF PERCENTAGE ('000)	ESTIMATED PERCENTAGE													
	0.1%	1.0%	2.0%	5.0%	10.0%	15.0%	20.0%	25.0%	30.0%	35.0%	40.0%	50.0%	70.0%	90.0%
1	171.4	170.7	169.8	167.2	162.7	158.1	153.4	148.5	143.5	138.3	132.9	121.3	93.9	54.2
2	121.2	120.7	120.1	118.2	115.1	111.8	108.5	105.0	101.5	97.8	93.9	85.8	66.4	38.4
3	99.0	98.5	98.0	96.5	93.9	91.3	88.6	85.8	82.8	79.8	76.7	70.0	54.2	31.3
4	85.7	85.3	84.9	83.6	81.4	79.1	76.7	74.3	71.8	69.1	66.4	60.6	47.0	27.1
5	76.7	76.3	75.9	74.8	72.8	70.7	68.6	66.4	64.2	61.8	59.4	54.2	42.0	24.3
6	70.0	69.7	69.3	68.2	66.4	64.6	62.6	60.6	58.6	56.5	54.2	49.5	38.4	22.1
7	64.8	64.5	64.2	63.2	61.5	59.8	58.0	56.1	54.2	52.3	50.2	45.8	35.5	20.5
8	********	60.3	60.0	59.1	57.5	55.9	54.2	52.5	50.7	48.9	47.0	42.9	33.2	19.2
9	********	56.9	56.6	55.7	54.2	52.7	51.1	49.5	47.8	46.1	44.3	40.4	31.3	18.1
10	********	54.0	53.7	52.9	51.5	50.0	48.5	47.0	45.4	43.7	42.0	38.4	29.7	17.2
11	********	51.5	51.2	50.4	49.1	47.7	46.3	44.8	43.3	41.7	40.1	36.6	28.3	16.4
12	********	49.3	49.0	48.3	47.0	45.6	44.3	42.9	41.4	39.9	38.4	35.0	27.1	15.7
13	********	47.3	47.1	46.4	45.1	43.9	42.5	41.2	39.8	38.4	36.8	33.6	26.1	15.0
14	********	45.6	45.4	44.7	43.5	42.3	41.0	39.7	38.4	37.0	35.5	32.4	25.1	14.5
15	********	44.1	43.8	43.2	42.0	40.8	39.6	38.4	37.1	35.7	34.3	31.3	24.3	14.0
16	********	42.7	42.4	41.8	40.7	39.5	38.4	37.1	35.9	34.6	33.2	30.3	23.5	13.6
17	********	41.4	41.2	40.5	39.5	38.4	37.2	36.0	34.8	33.5	32.2	29.4	22.8	13.2
18	********	40.2	40.0	39.4	38.4	37.3	36.2	35.0	33.8	32.6	31.3	28.6	22.1	12.8
19	********	39.2	39.0	38.4	37.3	36.3	35.2	34.1	32.9	31.7	30.5	27.8	21.6	12.4
20	********	38.2	38.0	37.4	36.4	35.4	34.3	33.2	32.1	30.9	29.7	27.1	21.0	12.1
21	********	37.2	37.1	36.5	35.5	34.5	33.5	32.4	31.3	30.2	29.0	26.5	20.5	11.8
22	********	36.4	36.2	35.6	34.7	33.7	32.7	31.7	30.6	29.5	28.3	25.9	20.0	11.6
23	********	35.6	35.4	34.9	33.9	33.0	32.0	31.0	29.9	28.8	27.7	25.3	19.6	11.3
24	********	34.8	34.7	34.1	33.2	32.3	31.3	30.3	29.3	28.2	27.1	24.8	19.2	11.1
25	********	34.1	34.0	33.4	32.5	31.6	30.7	29.7	28.7	27.7	26.6	24.3	18.8	10.8
30	********	31.2	31.0	30.5	29.7	28.9	28.0	27.1	26.2	25.2	24.3	22.1	17.2	9.9
35	********	28.8	28.7	28.3	27.5	26.7	25.9	25.1	24.3	23.4	22.5	20.5	15.9	9.2
40	********	27.0	26.8	26.4	25.7	25.0	24.3	23.5	22.7	21.9	21.0	19.2	14.9	8.6
45	********	25.4	25.3	24.9	24.3	23.6	22.9	22.1	21.4	20.6	19.8	18.1	14.0	8.1
50	********	24.1	24.0	23.6	23.0	22.4	21.7	21.0	20.3	19.6	18.8	17.2	13.3	7.7
55	********	23.0	22.9	22.5	21.9	21.3	20.7	20.0	19.3	18.6	17.9	16.4	12.7	7.3
60	********	22.0	21.9	21.6	21.0	20.4	19.8	19.2	18.5	17.9	17.2	15.7	12.1	7.0
65	********	21.2	21.1	20.7	20.2	19.6	19.0	18.4	17.8	17.2	16.5	15.0	11.7	6.7
70	********	20.4	20.3	20.0	19.4	18.9	18.3	17.8	17.2	16.5	15.9	14.5	11.2	6.5
75	****************		19.6	19.3	18.8	18.3	17.7	17.2	16.6	16.0	15.3	14.0	10.8	6.3
80	****************		19.0	18.7	18.2	17.7	17.2	16.6	16.0	15.5	14.9	13.6	10.5	6.1
85	****************		18.4	18.1	17.6	17.2	16.6	16.1	15.6	15.0	14.4	13.2	10.2	5.9
90	****************		17.9	17.6	17.2	16.7	16.2	15.7	15.1	14.6	14.0	12.8	9.9	5.7
95	****************		17.4	17.2	16.7	16.2	15.7	15.2	14.7	14.2	13.6	12.4	9.6	5.6
100	****************		17.0	16.7	16.3	15.8	15.3	14.9	14.4	13.8	13.3	12.1	9.4	5.4
125	****************		15.2	15.0	14.6	14.1	13.7	13.3	12.8	12.4	11.9	10.8	8.4	4.9
150	************************			13.6	13.3	12.9	12.5	12.1	11.7	11.3	10.8	9.9	7.7	4.4
200	************************			11.8	11.5	11.2	10.8	10.5	10.1	9.8	9.4	8.6	6.6	3.8
250	************************			10.6	10.3	10.0	9.7	9.4	9.1	8.7	8.4	7.7	5.9	3.4
300	************************			9.7	9.4	9.1	8.9	8.6	8.3	8.0	7.7	7.0	5.4	3.1
350	************************			8.9	8.7	8.5	8.2	7.9	7.7	7.4	7.1	6.5	5.0	2.9
400	********************************				8.1	7.9	7.7	7.4	7.2	6.9	6.6	6.1	4.7	2.7
450	********************************				7.7	7.5	7.2	7.0	6.8	6.5	6.3	5.7	4.4	2.6
500	********************************				7.3	7.1	6.9	6.6	6.4	6.2	5.9	5.4	4.2	2.4
750	**					5.8	5.6	5.4	5.2	5.0	4.9	4.4	3.4	2.0
1,000	**					5.0	4.9	4.7	4.5	4.4	4.2	3.8	3.0	1.7
1,500	**							3.8	3.7	3.6	3.4	3.1	2.4	1.4
2,000	**								3.2	3.1	3.0	2.7	2.1	1.2
3,000	**											2.2	1.7	1.0
4,000	**												1.5	0.9
5,000	**													0.8
6,000	**													0.7

NOTE: For correct usage of these tables, please refer to the microdata documentation.

Approximate Sampling Variability Tables - Northern Territories

NUMERATOR OF PERCENTAGE ('000)	ESTIMATED PERCENTAGE 0.1%	1.0%	2.0%	5.0%	10.0%	15.0%	20.0%	25.0%	30.0%	35.0%	40.0%	50.0%	70.0%	90.0%
1	****************		33.9	33.4	32.5	31.6	30.7	29.7	28.7	27.6	26.6	24.2	18.8	10.8
2	************************			23.6	23.0	22.3	21.7	21.0	20.3	19.5	18.8	17.1	13.3	7.7
3	************************			19.3	18.8	18.2	17.7	17.1	16.6	16.0	15.3	14.0	10.8	6.3
4	********************************				16.3	15.8	15.3	14.8	14.3	13.8	13.3	12.1	9.4	5.4
5	********************************				14.5	14.1	13.7	13.3	12.8	12.4	11.9	10.8	8.4	4.8
6	********************************				13.3	12.9	12.5	12.1	11.7	11.3	10.8	9.9	7.7	4.4
7	********************************				12.3	11.9	11.6	11.2	10.8	10.4	10.0	9.2	7.1	4.1
8	**					11.2	10.8	10.5	10.1	9.8	9.4	8.6	6.6	3.8
9	**					10.5	10.2	9.9	9.6	9.2	8.9	8.1	6.3	3.6
10	**					10.0	9.7	9.4	9.1	8.7	8.4	7.7	5.9	3.4
11	**					9.5	9.2	9.0	8.6	8.3	8.0	7.3	5.7	3.3
12	**						8.9	8.6	8.3	8.0	7.7	7.0	5.4	3.1
13	**						8.5	8.2	8.0	7.7	7.4	6.7	5.2	3.0
14	**						8.2	7.9	7.7	7.4	7.1	6.5	5.0	2.9
15	**						7.9	7.7	7.4	7.1	6.9	6.3	4.8	2.8
16	**							7.4	7.2	6.9	6.6	6.1	4.7	2.7
17	**							7.2	7.0	6.7	6.4	5.9	4.6	2.6
18	**							7.0	6.8	6.5	6.3	5.7	4.4	2.6
19	**							6.8	6.6	6.3	6.1	5.6	4.3	2.5
20	**								6.4	6.2	5.9	5.4	4.2	2.4
21	**								6.3	6.0	5.8	5.3	4.1	2.4
22	**								6.1	5.9	5.7	5.2	4.0	2.3
23	**								6.0	5.8	5.5	5.1	3.9	2.3
24	**									5.6	5.4	4.9	3.8	2.2
25	**									5.5	5.3	4.8	3.8	2.2
30	**										4.8	4.4	3.4	2.0
35	**											4.1	3.2	1.8
40	**												3.0	1.7
45	**												2.8	1.6
50	**												2.7	1.5
55	**													1.5
60	**													1.4
65	**													1.3

NOTE: For correct usage of these tables, please refer to the microdata documentation.

13 Contacts

TCESD Client Services
Client Services, Tourism and the Centre for Education Statistics Division
Statistics Canada,
150 Parkdale Avenue
Ottawa ON K1A 0T6
Tel: 1-800-307-3382
Fax : 613-951-4441
E-mail: educationstats@statcan.ca