Guide to Workforce Skills Assessment Instruments

Maria Fernanda Prada
Graciana Rucci

Guide to Workforce Skills Assessment Instruments

Maria Fernanda Prada
Graciana Rucci

IDB
Inter-American
Development Bank

Guide to Workforce Skills Assessment Instruments

Maria Fernanda Prada
Graciana Rucci
Labor Markets and Social Security Division
Inter-American Development Bank

Abstract

This guide reviews the principal instruments and surveys developed for assessing workforce skills. Its main objective is to consolidate the existing information and provide an overview of the different types of instruments available. It is also intended to help readers understand and navigate the vast universe of surveys and instruments for assessing skills in adults. For each instrument, it presents a general description, comparisons with other instruments of the same type, and an analysis of its applicability to the region.

JEL codes: J24, J0

Keywords: skills assessment, work skills, international surveys, skills measurement scales, skills demand surveys

I.   Introduction

Workforce skills and competencies are no doubt one of the greatest assets of any society. Skills are closely tied to earnings, increases in productivity, income distribution, economic growth, and well-being. (OECD, 1994, 2001; Hanushek and Woessmann, 2008; Heckman et al., 2006; Heckman et al., 2013, etc.).

The education that people acquire before entering the labor market, as well as the training they receive on the job, are determinants of human capital accumulation. However, direct measurements of investments in education and training are not necessarily a good proxy for workforce skills. Variables such as years of schooling or hours of training received are unlikely to provide accurate assessments of skills and competencies, especially in the presence of wide variations in quality and learning. Moreover, gaps in the standard variables (i.e., years of education and training) widen over time, because skills and competencies change rapidly and are increasingly complex. More in-depth knowledge about *how to assess work skills* is therefore essential.

It has also been shown that the diverse skills required for success in the labor market do not end with the competencies and knowledge traditionally assessed by standardized performance and academic achievement tests (Heckman et al. 2006; Borghans et al., 2008; Heckman and Kautz, 2013, etc.). Thus, prioritizing skills and competencies over years of training is what will ultimately impact the productive development of countries.

In this context, conducting these assessments and monitoring their results are important for several reasons: first, to determine country workforce potential; second, to determine the skills needed by the workforce–essential for developing strategies to boost the productivity of people who are already employed and for redesigning systems for training young people looking to enter the labor market with the skills needed by the productive sector; and finally, for monitoring the growth of human capital in countries and ascertaining the effects of policies designed to improve the skills of the workforce. Failure to do so entails very high costs that include lagging productivity, slow economic growth, limited efficacy of the resources allocated to education and training, underutilization of a country's human capital, and more.

In response to these demands, recent years have witnessed a rapid increase in the number of surveys and instruments for assessing workforce skills, including those of a socioemotional nature. These surveys and instruments differ in their content, objective, and methodology, focus on skills with similar names but different content, and respond to different needs.

This guide reviews the principal instruments and surveys developed for assessing workforce skills. Its main objective is to consolidate the existing information and provide an overview of the different types of instruments available. It is also intended to help readers understand and navigate the vast universe of surveys and instruments for assessing skills in adults. For each instrument, it presents a general

description, comparisons with other instruments of the same type, and an analysis of its applicability to the region.

A.   What are work skills?

1.   What do we mean by "skills"?

In the broadest sense, skills are the capabilities, competencies, qualities, talents, and in some cases, knowledge, that characterize individuals. In the context of the workforce, special emphasis is placed on the skill set or dimensions of skill that enable individuals perform successfully in the labor market.

It is essential, however, to recognize the significant differences in the terminology used when referring to skills. Heckman and Kautz (2013), for example, draw a fundamental difference between "skills" and "traits" that goes beyond semantics. In their view, traits connote permanence, immutability, and even genetic inheritance, while skills are abilities that can be changed – functional capacities that can be altered with interventions. For example, the term "skills" is often used in the context of certifying competency for specific jobs, referring in some cases to the capacity to perform certain tasks, and in others, personal qualities (attitudes, capacities) as well.

Other authors (Autor and Handel, 2013; Guvenen et al., 2015; etc.) reserve the term "ability" for the aptitudes an individual is born with and "skills" for the proficiencies necessary to perform work-related tasks. In this context, "ability" is directly related to individuals, whose portfolio of abilities is what enables them to learn the portfolio of skills required in the workplace.

In this guide, we make none of these distinctions and, on the contrary, have decided to use the term "skills" in its broadest sense, covering all dimensions measured by the instruments and surveys included in the review.

2.   Cognitive, socioemotional, and technical skills

There are numerous definitions and classifications of skills, since they vary with the county context, the very definition of "skill," the dimensions of skill considered, and many other factors. In this guide, however, skills are divided into three categories: cognitive, socioemotional, and technical. This classification is functional and broad enough to apply to the existing assessment instruments.

Following the convention used in the specialized literature, it identifies two major skills groups: *cognitive* and *socioemotional.*

Cognitive skills are associated with cognition and biological intelligence potential, or IQ (fluid intelligence), and acquired knowledge (crystallized intelligence).[1] This taxonomy can also be used to classify skills assessments.

The first group includes tests that assess skills considered necessary for acquiring knowledge and performing cognitive tasks; these include speed of perception, working memory, and inductive reasoning. These skills, and the tests that assess them, are highly specialized and are the object of research by cognitive psychologists, neuropsychologists, and other scientists.

The second group includes tests that assess more complex skills and involve integrating the components assessed in the first group and the knowledge acquired in specific areas. These skills are relatively easier to assess, because they are observed in the successful performance of specific tasks. This is true for the majority of standardized academic achievement tests that assess verbal or quantitative skills, vocabulary, general knowledge, etc. Table 1 presents this classification, together with some examples.[2]

Socioemotional skills, in contrast, are skills associated with an individual's personal qualities, attitudes, beliefs, personality traits, and behavior. They are also known in the literature as "soft skills," "behavioral skills," "21st century skills," and "social skills"[3] Psychology uses a fairly well-accepted taxonomy for these skills, known as the Big Five; this taxonomy describes the five most basic traits that can be used to characterize an individual and is used to classify other, more specific skills dimensions.[4] The Big Five are: Openness to Experience, Conscientiousness, Extraversion, Agreeableness, and Neuroticism, the latter of which negatively assesses emotional stability.

Two levels of socioemotional skills, analogous to the functional classification of cognitive skills, can be conceived: 1) basic personality skills or traits (Big Five and similar traits) and 2) skills that are a manifestation of a combination of these basic traits and can be directly seen in the workplace; they include communication, conscientiousness, cooperation etc. In the case of socioemotional skills, it is even harder to draw a bright line between these two groups, but we differentiate them for clarity's sake in the discussion. Table 1 presents this classification, together with some examples.

---

[1] This distinction is used mainly by psychologists and is known as the Cattell-Horn-Carroll theory (Carroll, 1993). Fluid intelligence is manifested through the acquisition of knowledge and experience, absorbing the culture and context in which one lives. Thus, the majority of the standardized tests mentioned and analyzed in this guide measure crystallized cognitive skills.

[2] It is useful to point out that the results of the exams and tests classified as cognitive can also measure non-cognitive skills. In fact, a large portion of the variations observed in the results of these tests can be attributed to personality and other factors such as incentive systems. See Kautz et al. (2014) and World Bank (2012).

[3] It should be noted that while these terms are sometimes used synonymously, they have different connotations. In this guide, we call these skills socioemotional, while recognizing the existence of alternative definitions. See Almlund et al. (2011) and Borghans et al. (2008) for a discussion and comparison of the different taxonomies.

[4] See Costa and McCrae (1992). See Table 1 in Kautz et al. (2014) for a detailed description of the Big Five personality traits, their definition according to the Dictionary of the American Psychological Association, highly correlated facets of personality, and the adjectives used to refer to these personality traits and some other related skills.

Finally, *technical skills* are defined in this guide as *specific* work skills and are thus directly relevant for a particular type of occupation or job. Their definition is therefore related to the characterization of jobs or occupations, the competency profile required for a certain type of job, and the description of the tasks performed in the different occupations.[5]

Again, the dividing line between technical and cognitive and socioemotional skills is blurry, since occupational profiles require a combination of skills in different dimensions. Notwithstanding, in this guide, we group them separately and recognize that their assessment is a response to different types of initiatives related much more to efforts to standardize occupations, classify jobs and vacancies, and determine employer recruitment needs.[6]

Table 1. Classification of skills and examples of work skills

Skills					
Types	**Cognitive**		**Socioemotional**		**Technical**
Subcategories	**Fluid**	**Crystallized**	**Basic**	**Manifestations**	
Examples of skills dimensions	Speed of perception, Working memory, Inductive reasoning	Language, Literacy, Numeracy, Problem solving	Big Five, Grit, Self-efficacy.	Problem solving, Communication, Conscientiousness, Leadership, Organization.	Computer skills, proficiency with specialized equipment
Examples of tests/assessment instruments	Raven matrices	Standardized tests. ALL, PIAAC, STEP, LAMP	Big Five, Rosenberg, Disconnected	Disconnected (employers), UKCESS	O*NET, STAMP[7], STEP, UKCESS
Location in this guide	Batteries of questions, questionnaires, and individual scales	International Skills Assessment Surveys	Batteries of questions, questionnaires, and individual scales	Skills usage/demand surveys	Skills usage/ demand surveys

Source: Prepared by the authors.

---

[5] These skills are generally acquired as part of regular work activities, professional development, on-the-job training, or experience interacting with others who possess these particular skills or competencies. (Pierre et al., 2014).

[6] See, for example, the initiatives for the creation of international classifications of occupations, U.S. initiatives such as the Dictionary of Occupational Titles (DOT) and its predecessor O*NET. A much more recent initiative is the Survey of Workplace Skills, Technology, and Management Practices (STAMP) (Handel, 2008), also found on O*NET, which measures the variety, intensity, and frequency of skills utilized in the workplace.

[7] This is a small study conducted in the United States, but it is interesting because it is one of the first surveys to measure workplace skills and requirements in terms of their content and the type of tasks required. This approach is known in the literature as the "task content approach," and several studies, such as STEP, have used it as an example or starting point. For more details, see Handel, 2008.

3. Definition of work skills

Definitions of the skill set considered "essential," "basic," or "fundamental" for the workforce abound. Some of them are merely functional, but most of them are the product of a consensus among experts forged in the particular circumstances of regions or countries. Consequently, they vary over time and from country to country and depend on the specific objectives for which they were created, as well as the institutions involved in their creation. These definitions include skills in both the three dimensions cited above – cognitive, socioemotional, and technical – and the various subcategories. Particular emphasis is placed on crystallized cognitive skills and manifestations of socioemotional skills because they are easier for employers, educators, and other stakeholders involved to identify and assess.

A classic example of this type of consensus-based definition is the skill set necessary for obtaining a job, identified in the early 1990s by the **Secretary´s Commission on Achieving Necessary Skills (SCANS)** in the United States. This commission drew up a list of the skills that a high school graduate should acquire for successful performance in the labor market.[8] This list, presented in Box 1, is often used as a reference in the literature.

---

[8] To arrive at this list, in addition to performing an exhaustive literature review, the commission consulted experts and interviewed workers and supervisors in 50 occupations.

Box 1: Necessary skills defined by SCANS

Basic competencies:

- Basic skills: Reading, writing, arithmetic and mathematics, speaking, and listening.
- Thinking skills: Thinking creatively, making decisions, solving problems, seeing things in the mind's eye, knowing how to learn, and reasoning.
- Personal qualities: Individual responsibility, self-esteem, sociability, self-management, integrity, and honesty.

Work competencies:

- Resource management: time, money, materials, space, and staff.
- Interpersonal relations: working on teams, teaching others, serving customers, leading, negotiating, and working well with people from culturally diverse backgrounds.
- Information management: acquiring and evaluating data, organizing and maintaining files, interpreting and communicating, and using computers to process information.
- Systems: understanding complex relationships, understanding systems, monitoring and correcting performance, and designing or improving systems.
- Technology: selecting equipment and tools, applying technology to specific tasks, maintaining and troubleshooting technologies.

Another example is the definition adopted by the OECD, used in its **PIAAC** survey to assess skills that are important "in many social contexts and work situations and necessary for fully integrating and participating in the labour market, education and training, and social and civic life." These skills include a combination of functional literacy (extensively defined), numeracy and language skills (reading, writing, digital literacy, and the proper use of mathematical concepts).

A common term is **"21st Century Skills."** Although it is used indiscriminately to refer to a variety of skill sets, a definition that is widely accepted and used is that of the Partnership for 21st Century Learning (P21), a U.S. organization founded in 2002 as a coalition of leaders in education, the business community, and other sectors charged with developing policies to determine the skills needed by the country's high school graduates (K-12) for successful performance in the labor market that are vital for the 21st century.[9] In this organization's view, student success in the new global economy requires a

---

[9] Some of the founding members of this coalition are: the U.S. Department of de Education, the AOL Time Warner Foundation, Apple Computer, Inc., Cable in the Classroom, Cisco Systems, Inc., Dell Computer Corporation, the Microsoft Corporation, the National Education Association, and SAP.

combination of knowledge, skills, and experience in five areas: 1) Knowledge in nine core academic subjects; 2) Interdisciplinary knowledge in areas such as global awareness; financial, economic, business, and entrepreneurial literacy; and civic, health, and environmental literacy; 3) Learning and innovation skills, which include creativity, critical thinking and problem solving, and, finally effective communication; 4) Information, media, and technology skills; and 5) Life and career skills.

Thus, life and career skills are but one dimension of the skill set and competencies that ensure success. These skills are a) flexibility and adaptability to change; b) initiative and self-direction to manage goals and time and work independently; c) social and cross-cultural skills to interact effectively with others and understand diversity; d) productivity and accountability to manage projects and produce results; and e) leadership and responsibility.[10]

These sample definitions show the wide variety of skills and competencies that the various assessments and surveys seek to capture. While they all have elements in common, it should be noted that skills can be interpreted and defined differently, depending on the context. An extensive list of definitions, along with a definition based on a review of more than 300 definitions, can be found in Lippman et al. (2015). Many of these definitions focus on important skills for children and youth.

Recently, within the framework of the "PRACTICE" model and employing a multidisciplinary approach, the World Bank compiled a list of socioemotional skills that are at the intersection between the skills valued by employers, those that predict academic and job success, those that follow an established development pattern, and, finally, those that can be changed through structured, evidence-based interventions. These skills are known as "skills for success."

The next section describes the content of this Guide to Workforce Skills Assessment Instruments.

B.   Content

This guide has two sections, an introduction and an inventory of surveys, tests, and instruments for adult skills assessment. The types of instruments and information provided for each are summarized below.

1.   Types of assessment instruments

---

[10] For a detailed description, see:

http://www.p21.org/storage/documents/docs/P21_Framework_Definitions_New_Logo_2015.pdf

The workforce skills assessment instruments are divided into three major categories: 1) International skills assessment surveys for individuals; 2) Batteries of questions or independent questionnaires commonly used to assess specific skills; and 3) Workplace skills usage and demand surveys for individuals or employers. Table 2 presents the assessment instruments contained in the inventory, arranged by the skill dimension they measure.

The ***International skills assessment surveys for individuals*** section includes an exhaustive list of international surveys of adults. The vast majority of these surveys assess an individual's skills, some include instruments for measuring skills usage in the workplace, and others supplement the survey of individuals with an employer survey. Six surveys are presented. The first five, **PIAAC, STEP, AHELO, ALL,** and **LAMP** are surveys conducted by various international organizations to measure the basic skills dimensions considered necessary for success in work and in life (aspects of literacy, numeracy, problem solving, and, in some cases, socioemotional skills).

This section also describes two surveys conducted in three Latin American countries under the name "**Disconnected**": the Skills and Trajectory Survey (STS) and an employer survey. These surveys are slightly different from the others but have been used extensively and have served as the basis for the creation or implementation of others in the region.[11]

The ***Batteries of questions, questionnaires, and individual scales*** section includes three initiatives for assessing specific skills dimensions. The list of these tests is endless and growing. This guide, therefore, mentions only initiatives that: 1) have made an effort to serve as a standard; 2) have rigorously selected the tests included; and 3) include low-cost tests that are available to the general public.

The first initiative is the complete battery of tests from the **IPIP**, which offers socioemotional skills and personality tests. The majority of these tests take their inspiration from the leading personality tests but are offered at no cost and are compared with other tests. The second initiative is the "**Toolbox" developed by the National Institutes of Health (NIH),** which contains an integrated package of tests that measure four skills dimensions: cognition, sensation, emotion, and motor. The tests have been validated for use with different cultures, ethnicities, geographic groups, ages, and types of studies. The Toolbox, moreover, is inexpensive and does not require the payment of royalties, uses state-of-the-art psychometric methodologies, and is intended to serve as a standard for medical and academic studies. The third initiative is the Social and Personal Competencies Scale (CPS), developed by the IDB, which measures socioemotional skills. The CPS is a brief, easy, and flexible test developed after a detailed review of the available instruments (See Brea, 2011).

These batteries of questions are useful to the extent that they are not tied to the administration of a general survey and can be used in the context of a specific project. The appendices contain a non-exhaustive list of commercial psychometric tests, as well as other instruments used to measure

---

[11] These surveys were conducted within the framework of an IDB project for the book of the same title (see Bassi, et al. 2012). Two IDB surveys are included under this title as part of the project: 1) the Skills and Trajectory Survey (STS), Chile and Argentina, and 2) the Employer Survey in Chile, Argentina, and Brazil.

specific skills components or dimensions. Table 2 contains the list of assessment instruments, arranged by the skills dimension they measure.

The last section presents ***skills usage and demand surveys***, which are particularly important due to their close connection with the world of work. In some cases, they are exclusively for businesses or employers and are used to analyze the necessary skills profiles; in others, they are for employees and are used to analyze the skills they need in their current jobs and to identify training needs and the skills that young people must have to enter the job market.

This guide includes three surveys that are representative of a large number of instruments: 1) The **UKCES Employer Skills Survey,** as a classic example of an employer survey used to identify the current and short-term demand for skills in the productive sector. [12] 2) The **Manpower-Talent Shortage Survey**, as an example of employer surveys in different countries that serve as measurements of perceptions in the productive sector, even though a detailed skills analysis is lacking. This survey was also chosen because it is often cited when discussing talent gaps. Given the widespread dissemination of its results, it is important to be familiar with its methodology and scope; and 3) The **European Working Conditions Survey (EWCS)**, as an example of an employee and employer survey deriving from the tradition of surveys that measure working conditions and have recently come to include the frequency and use of skills in the workplace.

Table 2. Assessment instruments by type of skill and section in the inventory

Skills					
Types	**Cognitive**		**Socioemotional**		**Technical**
Subcategories	**Fluid**	**Crystallized**	**Basic**	**Manifestations**	
1) International skills assessment surveys	Disconnected	PIAAC, STEP, AHELO, ALL, LAMP	STEP, Disconnected	CPS, Disconnected	STEP (households), Disconnected (employees)
2) Batteries of questions, questionnaires, and individual scales	NIH-Toolbox		IPIP, CPS		
3) Skills usage and demand surveys		STEP (employers), Disconnected (employers), STAMP		STEP (employers), Disconnected (employers), STAMP	STEP (employers), Disconnected (employers), STAMP

---

[12] Gonzalez-Velosa and Rucci (2016) describe current methods for anticipating skills demand. They present UKCESS as a special example of the employer survey method.

2.   Information on each instrument

The information on each instrument or survey is organized around the following 10 categories:

1) Organization responsible for its development and/or execution
2) General objective of the instrument,
3) History of the test, evolution, and background information
4) Target population
5) Type of skills assessed
6) Test content (modules, survey information, etc.)
7) Evaluation and interpretation of scores
8) Basic requirements for administration
9) Evaluation of its relevance and applicability to Latin America and the Caribbean
10) Administration of the test to measure the effects on variables of interest, such as success in the labor market, wages, etc.

C.   How to use this guide

This guide contains a compendium of information from different sources, presented in a comparable and detailed manner, and is meant to be used as a reference and not as a text to be read cover to cover. The intent is that after reading this introduction, researchers and policymakers will be able to determine which instruments will be useful to them and will skip directly to the Inventory section, where the instruments are described, to find additional information and useful references.

The next section of this introduction contains a summary and the main information about each of the three types of instrument included in the inventory.

Part 1, "International Skills Assessment Surveys for Individuals," provides an overview describing the nature, objectives, and comparability criteria for the surveys; a map indicating their geographic availability in the countries of the region; a list of general considerations to bear in mind when deciding whether to use them in a country; and finally, comparisons of the different surveys.

Part 2, "Batteries of questions, questionnaires, and individual scales," presents two tables summarizing the initiatives in the inventory, including the skills they assess, the organization responsible for their creation and/or execution, and their general characteristics. There is also a table with an expanded (but non-exhaustive) list of skills, indicating which ones are captured by the initiatives in the inventory and identifying other instruments commonly used to assess these skills. Finally, to assist the reader in

selecting batteries and questionnaires, a brief description of basic criteria is presented for comparison purposes, based on the psychometric properties of the batteries.

Finally, for the batteries in Part 3 of the inventory, "Skills usage/Demand," the summary and main messages section contains a table comparing the surveys included, based on the skills assessed by each.

For more details, for each battery in the inventory, there is a "Considerations for Users and Comments" section highlighting the most salient characteristics of each.

D.   Summary and main information

1.   International Skills Assessment Surveys for Individuals: PIAAC, STEP, AHELO, ALL, LAMP, and Disconnected

This section presents a summary and the main information about this type of international skills assessment survey in three dimensions. The first dimension, "Nature, objectives, and comparability," emphasizes the common elements of this group of surveys in terms of their nature, the objective for which they were designed, and broad comparability criteria. The second dimension, "Considerations," indicates general considerations to bear in mind when deciding whether to use one of these surveys in a country or to compare this type of instrument with others. The third dimension, "Comparisons among surveys," summarizes the main characteristics of the surveys and provides detailed comparisons of them.

Nature, objectives, and comparability

- These instruments are very similar. They are all part of a series of surveys that assess basic skills in adults and are highly comparable across countries. One of the first surveys of this type was the IALS (International Adult Literacy Survey) in 1994, now known as ALL (Adult Literacy and Lifeskills Survey). Subsequent surveys have since corrected the flaws of the initial ones, refining the definitions of "basic skills" and integrating new dimensions. See Table 1 for a comparison of the five surveys in terms of the skills dimensions they measure.

Table 3. International comparability of surveys by skills dimension

Dimensions/Topics	PIAAC	STEP	AHELO	ALL	LAMP
Cognitive					
Literacy/reading comprehension	X	X	X[4]	X	
Prose alone				X	
Document alone				X	
Reading components	X	X			X

Numeracy	X	x^2			X	x^3
Problem solving	X^1	x^2	X		X	
Critical thinking				X		
Analytical reasoning				X		
Socioemotional			X			
Technical and specific skills						
Use of skills in the workplace	X	X				

Source: Prepared by the authors, using survey data. Notes: (^{1}) Assesses problem solving in technology-rich environments. (^{2}) Self-reporting. No direct assessment. (^{3}) Assesses only use of numbers. (^{4}) Assesses written communication.

- Given the nature of these surveys, it is useful to compare them according to the dimensions they assess, their geographic coverage, and their capacity to accommodate to the specific needs of a country. Figure 1 shows the geographic availability of these surveys in the countries of the region.

Figure 1. Geographic coverage of surveys in the region
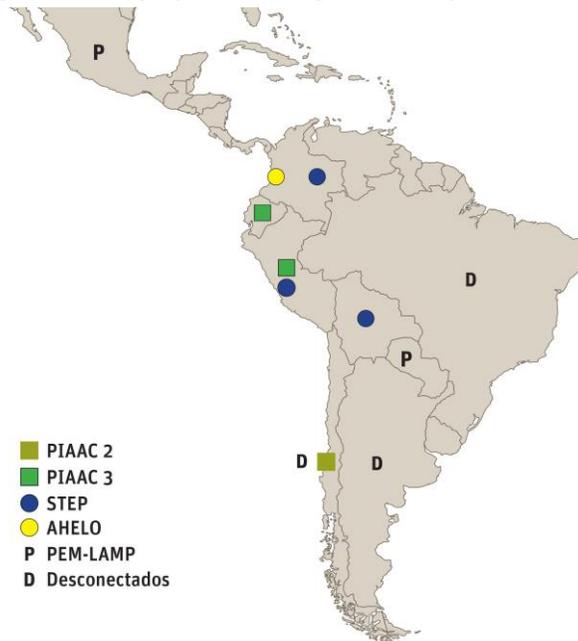
Source: Prepared by authors, using information published on the respective assessment websites.

Notes: PIAAC 2 and 3 refer to the second and third round, (2012-2016) and (2016-2019), respectively. Mexico could be part of PIAAC 3. Desconectados refers to the book "Disconnected"

- Country participation in this type of initiative is always rewarding, since it is an opportunity to compare results with those of other countries, even with respect to a few skills that are not necessarily high in demand in the productive sector. This type of exercise is also important for the countries of the region interested in joining the OECD, since it builds bridges of communication between the organization and the countries and generates useful information for country studies that are comparable with those of the member countries.
- All the surveys in this category should be considered as country projects. The skills assessments cannot be implemented in isolation of the whole project. So if the reader is looking for a battery to measure specific skills, she should refer to the section "Batteries of questions, questionnaires, and individual scales" in this document.

Comparisons of specific surveys

- PIAAC can be considered the most recent version of the surveys that assess literacy, numeracy, and problem solving–three skills basic to success in work and life (OECD, 2012). Like ALL, PIAAC assesses numeracy, while IALS assesses a similar, but more basic, concept (numerical literacy). PIAAC is the only survey that assesses literacy and problem solving in technology-rich environments (ALL measures problem solving, but not in technology-rich environments).
- A serious disadvantage of PIAAC is that, at least to date, it does not assess socioemotional skills, which limits the range of skills that are comparable among countries.
- The STEP survey has one advantage over all the others, which is that it includes assessments of socioemotional skills. Furthermore, it is designed for low-to-middle-income countries and thus has modules that are relevant for those countries (for example, household information, modules for assessing socioemotional skills tailored to populations with high levels of illiteracy, etc.).
- LAMP is not radically different from other, more recent and complete surveys like PIAAC and STEP. One of these two would therefore be preferable for assessing skills for work and life.
- Unlike PIAAC and IALS, LAMP is more global in nature and is focused on developing countries with a greater number of linguistic families and writing systems. This means that it is better suited to the needs of countries whose populations have low levels of education, since it is simpler in terms of content and administration.
- An important consideration is the fact that LAMP was designed to contribute to national capacity building, drawing from existing experience and promoting it as the best mechanism for supporting national initiatives. This, in turn, means that it encourages countries to improve their institutional environment (important for data collection, training, and institutional work).
- AHELO is not strictly comparable to the other international surveys reviewed, because up to now it has been used to determine the feasibility of its systematic implementation in several countries. It is included in the inventory because it is the only survey focused on evaluating the

quality of higher education and determining whether it is possible to calculate how much of educational achievement can be attributed to the institutions.

- AHELO compares institutions in terms of the level of preparation of graduates on completing their studies. It compares the programs in each institution and measurements of value added, based on variables of interest. Its ultimate objective is to provide information to increase the efficiency of the institutions.
- "Disconnected" is a survey developed in response to the lack of information about skill mismatch and skills assessment in the region. As of this time, there are no plans to repeat it for other countries or periods.
- It should be underscored that in designing the Disconnected survey, a real effort was made to select skills that can easily be related to the type of skills demanded in the productive sector, thus building a conceptual bridge between the skills studied in academia and perceptions of what is required in the workplace.
- Unlike the international surveys conducted by organizations such as the OECD, the World Bank, UNESCO, etc., Disconnected's skills assessment modules are available for uses other than those initially envisaged for the survey (once the access and usage protocol has been completed).

Table 4. Comparison of the most recent surveys in relevant dimensions

Dimensions	PIAAC	STEP
Use of skills by:	Workers	Workers and the employers for whom they work (including **informal**)
Time:	90-115 min	120-150 individuals/45-60 employers
Number of countries:	Round 1: 24; Round 2: 33 (Chile), Round 3: 39 (Ecuador, Mexico, Peru)	10 (Bolivia and Colombia)
	Focus on developed countries. Includes OECD members and other countries (Latin American countries in parentheses)	Focus on developing countries
Approximate cost:	Cost varies from country to country: exceeds €1 million plus Fixed annual cost of €100,000 (for countries that are not OECD members)	US$470,000-530,000 (3,500 households).
		US$111,000 (400 employers)
Length of the cycle:	56 months (19 for implementation)	28 months for data gathering
Response rate:	45-73%	43-98% individuals/38-51% employers
Sample:	4,500 and 27,399 individuals	29,00-4,000 households/300-400 employers
Representativeness:	National	Urban

		Self-reporting of usage by individuals vs. employer demand
Skill mismatch:	Skills measured vs. self-reporting of skills usage	Skills assessed (reading and socioemotional) vs. those required by employers

2. Batteries of questions, questionnaires, and individual scales: IPIP, NIH Toolbox, and CPS

This inventory includes only three initiatives for measuring the specific skills dimensions: the **"International Personality Item Pool" (IPIP)**, developed by the University of Oregon; the "**Toolbox,**" developed by the National Institutes of Health (**NIH**); and the **Social and Personal Competencies Scale (CPS)**, developed by the IDB.[13] These initiatives were selected because they possess three characteristics: 1) they were designed as standards; 2) the tests included were subjected to a rigorous selection process; and 3) they include tests that the general public can access at little or no cost.

Table 5, below, summarizes the initiatives.

Table 5. Initiatives included in the inventory and main characteristics

Initiative	Skills	Individual/ institution in charge	Characteristics
IPIP	Socioemotional and personality	Lewis R. Goldberg, University of Oregon (Oregon Research Institute)	- Continuous improvement initiative - Easy access and help for users - Alternative to the main personality tests - Used and validated in some Latin American countries
NIH Toolbox	Cognitive, sensation, emotional, and motor	National Institutes of Health (NIH) Attn: Cindy Nowinski, MD PhD 312-503-4800	- Tests validated in several cultures, ethnic groups, ages, and types of study - Uses state-of-the-art psychometric methods - Designed to serve as the standard for medical and academic studies - No evidence of use in Latin America at this time. - Not yet used in economic contexts or the labor market. Rather than a disadvantage, this is an opportunity for initiatives in the region. - Free access to programs funded by NIH, but low-cost access can be requested for other projects.

[13] Duryea and Williamson, 2016 contains a detailed description of various tests and indicators, including the CPS, for assessing socioemotional skills (psychosocial competencies) in youth.

CPS	Socioemotional	IDB	• Short, flexible, and easy to administer. • Designed after a detailed review of the available instruments. • Used and validated by several countries in the region for different populations. (Ibarrarán et al., 2012)

These batteries of questions are useful because they are not tied to the administration of a general survey and can be used in the context of a specific project. As a reference, Tables A1 to A3 of the Appendix present a non-exhaustive list of aptitude tests offered by employers, as well as several other scales, batteries, inventories, and instruments used to measure specific skill components or dimensions.

For a general overview and discernment when selecting or reviewing these scales, Table 6 lists an extensive (non-exhaustive) set of skills, indicating which are captured by the initiatives included in the inventory, in some cases including the names of the tests most commonly used to assess these skills.

Table 6. Initiatives included in the inventory, main skills, and commonly used tests

Skills	Associated facets and skills	IPIP	NIH Toolbox	CPS	Other commonly used tests
Socioemotional skills					
Openness to Experience	Imagination	x			NEO-PI-R, 16PF, AB5C, CPI, MPQ, 6PFQ, VIA, AMBI
	Creativity	x			Torrance Tests of Creative Thinking (TTCT), Structure of Intellect (SOI) divergent production plane
Conscientiousness	Conscientiousness	x			NEO-PI-R, 16PF, AB5C, CPI, VIA, AMBI
	Competence	x			
	Order	x		x	
	Goal orientation / Ambition	x			
	Grit/perseverance/ determination	x			Duckworth (2009)
Extraversion	Extraversion				Adult Eysenck Personality Questionnaire (AEPQ), NEO-PI-R, 16PF, AB5C, AMBI,
	Warmth	x			
	Assertiveness	x			
	Ability to relate	x	x		Prosocial Behavior Scale
Agreeableness	Altruism	x			
	Empathy-Ability to Communicate	x		x	Empathy Quotient test (EQ), Index of Empathy for Children (Bryant)
	Cooperation	x			
	Competitiveness	x			
Neuroticism	Anxiety	x			

	Hostility (irritability)	x			
	Depression	x			
	Impulsiveness	x			Chan and Henry, 2009 (aggressivity), go/no-go activity (inhibitory control, impulsiveness, and inattention). What would make you fight activity (aggressivity)
	Emotional well-being		x		
	Vulnerability to stress	x	x		
	Negativity		x		
	Locus of control (internal vs. external)	x			Rotter (1966), Tangney (self-control)
	Self-esteem	x		x	Rosenberg (1965)/ Albo et al. (2007) in Spanish, Coopersmith Self Esteem inventory
	Self-efficacy		x		MSQL, Students Approaches to Learning (SAL), Disconnected, General Self-efficacy Scale (GSE)
Other Big Five tests					Single-Item Measure of Personality-SIMP (Wood and Hampson, 2005); NEO-PI-R, 16PF, AB5C, AMBI, TIPI Scale (Gosling et al., 2003)[14], Rammstedt and John 10-item Scale (2007)
Leadership				x	Leadership Skill Inventory, Life Skills Evaluation System
Conflict behavior				x	
Empathy and Ability to communicate				x	
Cognitive and other skills					
Cognitive	Executive function, attention, episodic memory, language, processing speed, and working memory		x		Flanker (attention and executive function), WASI- Wechsler Abbreviated Scale of intelligence
Motor	Dexterity, strength, balance, gait, and endurance		x		
Sensation	Hearing, vision, vestibular capacity, smell, taste, and pain		x		

Source: Prepared by authors, adapting the table of John and Srivastava (1999) in order to group socioemotional skills.

a) *Selection criteria*

The first step in selecting the right test is to consider and evaluate its psychometric properties to ensure that the instrument measures what it is supposed to (validity), that the results can be trusted

---

[14] Ten Item Personality Measure (TIPI) Gosling, S. D., Rentfrow, P. J., and Swann, W. B., Jr. (2003). Renau et al. (2013) present the Spanish version, validated by a sample in Spain.

(confidence), and, if a standard for interpretation is needed, that a standard has been constructed for a population similar to the one to which the test will be administered.

A brief description of each of these properties is presented below. The most important of these are validity and confidence. Validity means that the test measures what it was designed to measure and that the results and inferences are valid.[15]

a.1) Reliability

- Definitions:
  - Confidence means that the answers are stable and consistence across different test measures. This ensures that the results are replicable.
  - In psychometrics, confidence refers to the general consistency of a measurement. Thus, a measurement is said to be consistent if it yields similar and stable results under different measurement conditions.

- Measurements

The simplest way to measure the confidence of an instrument/test or measurement is to take the ratio between the theoretical variance (for example, the latent or innate skill) and the observed variance (the measurement variance, which includes other confounding factors). This ratio can yield values ranging from 0 and 1, where 0 represents no reliability and 1 represents perfect confidence. This ratio is known as the reliability coefficient or confidence level.

$$\rho = \frac{\sigma_t^2}{\sigma_o^2} = 1 - \frac{\sigma_e^2}{\sigma_o^2}$$

where ρ is the confidence coefficient, σ_t^2 is the theoretical (true) variance, σ_o^2 the observed variance, and σ_e^2 the error variance. Since the theoretical variance is unknown and cannot be measured directly, a series of methods can be used to estimate it.

The best-known of these is the Cronbach's alpha coefficient, or Cronbach's alpha (α) (Cronbach, 1951), which is actually an indicator of internal consistency and indicates the degree to which the test items co-vary. As a general criterion, George and Mallery (2003, p. 231) suggest the following for evaluating Cronbach's alpha: 0.9 is excellent, 0.9-0.8 is good, 0.8-**0.7** is **acceptable**, 0.7-0.6 is questionable, 0.6-0.5 is poor, and less than 0.5 is unacceptable. While there is some debate about this, the highest expected value is 0.90, since anything above this value implies redundancy or duplication. (Streiner, 2003).

- Related concepts:

  - Internal consistency: Measures whether the items or questions in the scale measure the same concept. Measures the consistency of the results among the different test items. Apart from Cronbach's alpha, some of the common methods for measuring internal consistency are the Kuder-Richardson coefficient, Rulon's Method, the Guttman/Flanagan Method, the beta coefficient (β) and the theta (θ) and omega (Ω) coefficients.

---

[15] This discussion follows the parameters of the psychometrics appendix of Wilson-Ahlstrom et al (2014).

- o Test-retest Reliability: Measures the consistency of test results at different points in an evaluation, the stability of the results over time. This coefficient is relevant only when the measurements are made in a brief period of time when major changes in the aspect to be measured are not anticipated (for example, an individual's skills as part of his natural development or participation in a program).
- o Interrater Reliability**:** Measures the agreement between two or more evaluators. It is especially important for tests requiring an external evaluator (this is not the case for self-reported tests). In this case, simple measurements, such as percentages of agreement, kappa statistic, or intraclass correlation coefficient are used.
- o Consistency between methods: Measures the consistency of the results with variations in the measurement methods or instruments used.

- Key questions

  - o Is there evidence that the instrument's scales yield consistent response?

  - o How robust is the evidence of the test's reliability? This refers to the Cronbach's alpha value.

  - o To which population groups does the evidence of confidence apply? Evidence of confidence for one group does not guarantee the test's confidence for other groups. Efforts should be made to ensure the use of tests that are reliable for the group of interest or a similar group; if the tests are not, this factor should be considered when deciding whether to use them.

a.2) Validity

- Definitions:
  - o Validity refers to the ability of the instrument/test to measure what it was intended to.

  - o Validity is a concept that refers to the ability of a measurement instrument to accurately measure the trait it claims to measure. Thus, a measurement instrument is valid insofar as empirical the evidence legitimizes the interpretation of the test scores. (Wikipedia).

- Measurement:

  - o Unlike confidence, validity is hard to measure, since the relevant concepts are intangible. Thus, validity is more a qualitative consideration, and there is no number to compare with others.

  - o Several types of statistical analysis are used to inform judgments of validity, some of which are discussed below.

  - o Construct validity: Measures how related the test results are to the measurements of other concepts theoretically associated with the psychological construct in question. Operationally, factor analyses are used; these are multivariate analysis methods that examine the correlation between different questions or scales to find the common measurement element.

o Predictive validity: Also known as "criterion validity," "criterion-related validity," or "prognostic validity," this refers to the effectiveness with which a variable of interest (criterion) can be predicted or forecasted based on the test results (correlations with the variable of interest).

o Convergent validity: Compares two similar tests that measure the same concepts and calculates how great the association between their results is.

o Discriminant validity: Compares two tests constructed to measure different concepts and calculates how small the association between their results is.

- Key questions

    o Is there evidence that the scales on the instrument measure what they intend to measure, that is, are valid? Generally, the question is whether serious studies have been published containing statistical analyses of validity using factor analysis and the methods mentioned in the preceding section.

    o How robust is the evidence? What methods are used and what are the results?

- Considerations:

    o Reliability is a necessary, but insufficient, condition for guaranteeing validity. A test can be reliable and consistently measure a construct that is different from the construct of interest. For example, there are a number of consistent tests of specific skills, but not all of them are valid for predicting job performance.

    o Although validity does not imply confidence, a test's lack of confidence calls its validity into question.

3. Skills usage/demand surveys

The last section of this guide presents three surveys representative of a large number of *surveys that measure skills usage and demand.* These surveys are relevant because they measure skills from the employer's standpoint and use skills definitions and measurements focused on specific occupations or jobs. This is a fundamental difference with the surveys in the first section, in which skills measurement is not necessarily focused on employment and the workplace and terminology and concepts may be used that are unfamiliar to employers or the employees themselves.

In this type of study, the unit of analysis is not always the same. Some studies are geared exclusively to businesses or employers and analyze the skills profiles they need, while others concentrate on employees and are used to analyze the skills needed in their current job and to identify training needs and the skills needed by young adults to enter the labor market.

This first survey of this type in the guide is the employer survey developed by the UK Commission for Employment and Skills, known as the **"UKCES Employer Skills Survey."** This survey is presented as a classic example of a <u>national employer</u> survey designed to determine current and short-term skills demand in the productive sector.[16]

The second survey is the ManpowerGroup **"Talent Shortage Survey,"** included as an example of employer surveys in <u>different countries</u> that, despite the lack of detailed skills analysis, serve as measurements of perceptions in the productive sector. It was also chosen because it is often cited in discussions of <u>talent gaps</u>. Given the extensive dissemination of its results, it is important to be familiar with its methodology and scope.

The third survey is the **European Working Conditions Survey (EWCS)**. This was included as an example of an employee-employer survey that hails from the tradition of surveys designed to measure working conditions and have recently evolved to include the frequency and use of skills in the workplace.

Table 7 summarizes the surveys contained in this guide by type of skills assessed.

Table 7. Comparison of skills usage and demand surveys

Skill	UKCES Employer Skills Survey	Manpower Talent Shortage Survey 2014	EWCS	Disconnected
Cognitive				
Reading and writing skills				X
Numerical literacy				
Numeracy	X			X
Computer skills	X		use	use
Advanced IT or software skills	X			
Socioemotional				
Generic socioemotional skills		X		
Language and communication (oral and written)	X			X
Problem solving	X			X
Critical thinking				
Analytical reasoning				
Customer relations	X		use	X
Collaboration	X			

---

[16] Gonzalez-Velosa and Rucci (2016) describe the different methods used to predict skills demand. The UKCES is presented as a particular example of the method based on employer surveys.

Planning and organization	X			
Strategic management	X			
Positive attitudes and behaviors				X
Conscientiousness				X
Technical				
Technical, practical, or specific generic skills	X	X		
Willingness to learn the specific skills required				X
Use of specific machinery			X	
Manual skills			X	
• Exposure to vibration from manual tools or machinery			X	
• Repetitive hand or arm movement			X	
• Lifting heavy weights			X	
• Others related to occupational risks			X	

Source: Prepared by authors, using official survey data.

Skills usage and demand surveys vary widely and are generally used for purposes beyond skills measurement. A good description of employer skills surveys can be found in Hogart (2016). In addition, readers interested in studying employer surveys to anticipate the demand of employers and the productive sector in general should consult Gonzalez-Velosa (2016).

II. INVENTORY

Part 1: International Skills Assessment Surveys for Individuals

This section presents international surveys for adults. The vast majority of these surveys assess the skills of individuals, some include instruments that measure skills usage in the workplace, and others supplement the survey of individuals with surveys of employers.

1. PIAAC: Survey of Adult Skills (Programme for the International Assessment of Adult Competencies, International Survey of Adult Skills

1.1. Organization responsible for development and/or execution

OECD

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

1.2. General Objective

"The OECD Programme for the International Assessment of Adult Competencies (PIAAC) is an initiative of the OECD that assists governments in assessing, monitoring, and analysing the level and distribution of skills among their adult populations as well as the extent of skills use in different contexts. The Survey of Adult Skills is a central pillar of PIAAC, providing an unparalleled source of evidence for policy makers on the skills of adults."[17]

This survey assesses the basic competencies of adults aged 16 and onward that are "relevant for adults in many social contexts and work situations, and necessary for fully integrating and participating in the labour market, education and training, and social and civic life."

The survey assesses skills in three primary domains (sublevels): literacy, numeracy, and problem solving in technology-rich environments.

Moreover, it provides information on the participants' reading- and math-related activities and use of information and communication technologies in the workplace and daily life. It also provides information on a series of generic skills that are essential in the workplace, such as cooperation and time

---

[17] http://www.oecd.org/skills/piaac/Key%20facts%20about%20the%20Survey%20of%20Adult%20Skills.pdf

management. Finally, it asks respondents if these skills and qualifications are what they need for their job and whether they can act independently in key areas of their work.

1.3. History

Round 1 (2008-2013):

24 countries: 22 OECD (Australia, Austria, Belgium (Flanders), Canada, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Ireland, Italy, Japan, Korea, Netherlands, Norway, Poland, Russia, Spain, Slovak Republic, Sweden, the United Kingdom (only England and Northern Ireland), and the United States.

Round 2 (2012-2016):

33 countries: 24 countries that participated in round 1 plus Chile, Greece, Indonesia, Israel, Lithuania, New Zealand, Singapore, Slovenia, and Turkey.

Round 3 (2018-2023) or (2014-2018), depending on the source[18]:

39 countries. According to the OECD, Mexico, Ecuador, and Peru are in the data collection phase.

PIAAC is the third in a series of adult skills surveys that began with IALS (1994-98) and ALL (2003-06). The three surveys assess prose and document literacy combined. Unlike IALS and ALL, PIAAC does not assess these components separately. Like ALL, PIAAC assesses numeracy, while IALS assesses a similar, but more basic concept quantitative literacy. PIAAC is the only survey that assesses literacy and problem solving in technology-rich environments (ALL assesses problem solving but not in technology-rich environments).

PIAAC has made improvements in at least three methodological aspects: greater country coverage (33 countries vs. 11 and 15, respectively); larger sample sizes per country (almost double those of IALS), which facilitates comparisons among groups; and a greater number of skills assessed.

1.4. Target population
Adults aged 16-65 residing in the country at the time of the survey, regardless of their nationality, citizenship, or language.

1.5. Type of skills assessed

---

[18] Check the current status of round three (dates and countries) in the original source
http://www.oecd.org/skills/piaac/aboutpiaac.htm

This survey assesses skills in three main domains: literacy, numeracy, and problem solving in technology-rich environments. Literacy is assessed with a general test; this is supplemented with an additional section on reading components that assesses basic skills related to or dependent on literacy. Proficiency in mathematics is assessed basically through numeracy.

The survey also provides information about a set of generic skills that are essential in the workplace, such as cooperation and time management.

Literacy

Literacy is the ability to understand and use written information in a variety of contexts to achieve different goals and increase the knowledge and potential of individuals. It is one of the basic requisites for developing higher-order skills and obtaining positive economic and social outcomes. Previous studies have shown that literacy is closely associated with positive work outcomes, as well as social engagement and life-long learning.

Unlike earlier literacy assessments, this study assesses the ability of adults to read digital texts (for example, texts containing hypertext and navigation features, such as scrolling or clicking on links), as well as printed texts.

It assesses three cognitive processes: access and identification, interpretation and integration, and evaluation and reflection.

Reading components

For more detailed information on adults with low reading levels, the literacy assessment is supplemented with a test of reading components, which measure the basic decodification skills that enable people to extract meaning from written texts: knowledge of vocabulary, the ability to process meaning at the sentence level, and speed in reading passages of text.

Numeracy

Numeracy is the ability to use, apply, interpret, and communicate mathematical information and ideas. It is a basic skill in an era when individuals are confronted with growing amounts of quantitative and mathematical information in daily life. It is also a parallel competency to literacy, and it is important to determine how these two competencies interact, since they are distributed differently in different subgroups of the population.

Solving problems in technology-rich environments

This refers to the ability to use technology to solve problems and perform complex tasks. It is not an assessment of "computer literacy" but of the cognitive skills needed in the Information Age, in which the unlimited accessibility of information has made it essential for people to be able to decide what information they need, critically evaluate it, and use it to solve problems. In this evaluation, both higher-order and more basic skills are assessed.

Generic skills – work competencies

The survey asks adults about the intensity and frequency with which they use these skills in the workplace. These are not direct skills tests, but questions of an informational nature that are part of the background questionnaire.[19]

This information covers four major categories: cognitive skills, interaction and social skills, physical skills, and learning skills.

Cognitive skills consist of reading, writing, mathematics, and the use of information and communication technologies.

Interaction and social skills consist of collaboration and cooperation, planning and the use of one's own or others' time, communication and negotiation, and contact with customers.

Physical skills involve the use of motor skills.

Learning skills involve activities such as instructing others, learning (formally or informally), and keeping up-to-date in one's professional field. In addition, all participants are asked about the frequency and intensity of their reading, writing, and numeracy activities, as well as their use of information and communication technologies in the home and other settings.

1.6. Content:

The survey contains a background questionnaire and a skills assessment module.

Background questionnaire: The test contains a background questionnaire that collects a variety of information about the factors that influence skills development and maintenance, such as education, experience, formal and informal training in the 12 months prior to the survey, social background, interaction with ICTs, language, and other types of skills-related information (skills usage in the workplace). The information paints a picture of the respondents' current situation, employment status, and income. In addition to economic aspects, the evaluation includes questions about health, social participation, political efficacy, and social trust.

---

[19] PIACC took its inspiration from the Job Requirements Approach (JRA), first implemented in skills surveys in the United Kingdom (UK Skills Survey) (Felstead et al., 2007).

Skills assessment module: The content of this module varies with the format used. The survey is administered on paper to individuals with no previous computer experience and to those who, despite their previous experience, do not receive the minimum score on the ICT assessment test.

Computer-based assessment:

- Test to assess information and communication technologies (ICT): Involves basic tasks to assess the minimum skills for using a computer.

- Basic literacy and numeracy skills: a series of short, simple questions and tasks.

- Literacy: 52 questions, 30 of them taken from the IALS questionnaires to permit some degree of comparison. Measures three basic cognitive processes: access and identification, interpretation and integration, and evaluation.
- Numeracy: 52 questions, 30 of them taken from the IALS questionnaires to permit some degree of comparison. Measures five basic cognitive processes: identify, locate/access; act upon, use; interpret, evaluate or solve; and communicate.
- Ability to solve problems in technology-rich environments: 14 questions created exclusively for this test. Assesses the ability to perform tasks that include several steps and, in some cases, the use of assorted technologies (e-mail, spreadsheets). Assesses four basic cognitive processes: setting goals and monitoring progress; planning; acquiring and evaluating information; and using information.

Paper-based assessment

- Basic assessment of literacy and numeracy skills: a series of short, simple questions and tasks.

- Literacy: 24 questions, all of them taken from the IALS questionnaires; 18 of these are the same as those contained in the computer-based version.
- Numeracy: 25 questions taken from IALS questionnaires; 21 of these are the same as those contained in the computer-based version.
- Reading components: reading vocabulary (34 questions), understanding the literal meaning of sentences (22 questions) and comprehending multi-paragraph passages (4 questions). Designed to assess the skills of the population with low skill levels and to capture information on times and accuracy. Print vocabulary (participant is shown an image and must select the correct word out of four options). Sentence processing (participant must make a judgement about the accuracy of a sentence). Passage comprehension (participant must complete the written text by selecting the correct word from a group of options).

1.7. Evaluation–scoring and interpretation

Competencies are assessed using a 500-point scale, divided into levels. Each level summarizes what a person is capable of doing with a certain score. There are five levels in literacy and numeracy and three in problem solving.

[Table in appendix]

1.8. Basic requirements for administration

The country must agree to administer the two required components, literacy and numeracy.

The sample sizes depend on the particular interests of the countries. In principle, the survey must be nationally representative (samples range from 4,500 to 27,300).

The survey must be administered under the supervision of trained interviewers. The interviewer administers a background questionnaire. **Time: 30-45 minutes** (20 minutes of basic questions plus one of three 10-minute variants).

After the background questionnaire, participants unfamiliar with computers hand in the paper version and the rest of the adults take the computer-based skills tests. These tests have no time limit but take **60 minutes** on average. (For the paper-based test: 10 minutes of diagnostic questions; if a minimum number of questions are answered correctly, the participant is randomly assigned 30 minutes of literacy and numeracy questions, plus 20 additional minutes of component skills. Participants who fail to correctly answer a minimum number of diagnostic questions must answer the reading component questions. For the computer-based test: 10 minutes of diagnostic questions in two stages; if a minimum number of questions are answered incorrectly in the first stage, the individual is given the paper test. If a certain number of questions are answered incorrectly in the second stage, which contains six cognitive questions, the participant must complete only the reading components section. The rest of the respondents are administered one of three questionnaires, 50% of which assess a combination of literacy and numeracy skills; 33%, a combination of problem solving with reading or language skills; and the remaining 17%, only problem solving).

1.9. Relevance for Latin America

High comparability among countries and over time.

Criteria for country participation in the survey: 1) follow OECD guidelines for administering the survey (PIAAC Technical Standards and Guidelines); 2) guarantee the confidentiality of materials related to the survey; 3) comply with the agreed dates for publishing the results. 4) furnish the information to the participant, the OECD, and the international contractor; 5) annual cost of €**100,000.** If the country decides to discontinue its participation, it must send notification 12 months in advance and pay a fine. (€10,000 per year).

Country responsibilities: 1) abide by the technical standards of the test; (2) tailor the questionnaire to the context; 3) translate the questionnaires and instruments after receiving training from the PIAAC group; 4) collaborate in verification of the questionnaires and survey; 5) adapt the PIAAC training materials and conduct the national training; 6) process the information and create national databases using the PIAAC software, including national reports and results; 7) appoint at least four delegates: the country project director, the sampling director, the IT coordinator, and the national data director; 8) cover the cost of administering the survey (this could come to €1 million, depending on the country).

Each round of the survey takes approximately 56 months from the opening of talks to dissemination of the results. The process consists of four phases: preparation of the pretest (14 months), preparations for administration of the pretest (8 months, including 2 months of training), preparations for the administration of the survey (**19 months**, including 2 months of training, and 11 months for data collection).

1.10.   Administration

1.11.

PIAAC is an initiative designed to influence public policies and the design and planning of social interventions, as well as education and training programs. More specifically, it provides information about the situation and dimensions of illiteracy or low literacy levels, the mismatch between demand in the labor market and skills training in the education and training systems, equity in access to education and intergenerational mobility, the transition of young adults to their first job, the identification of at-risk populations, the association between cognitive skills scores and demographic variables, the education and heath context, etc.

- Correlation between scores and years of education: Seven additional points in numeracy are equivalent to one year of education (education-numeracy correlation coefficient=0.45), using the average for the 24 countries that administered the survey in the first round. (Hanushek et al 2013).

- Correlation between scores and the probability of being employed, high wages, health, and other social behaviors.

  - Hanushek et al (2013) measure the returns of cognitive skills for adults aged 35-54 (using the natural logarithm of wages per hour). Controlling for gender and work experience, an increase of one standard deviation in the numeracy score increases wages by 18%.

- Correlation between the average score obtained by countries and growth, productivity scores, inequality. (OECD, 2013).

- Estimates of changes in skills over time, estimates of the age-skills profile of cohorts over time and in the life cycle for each cohort (using PIAAC and ALL and IALS estimates). Desjardins, R. and A. Warnke (2012) compare literacy over time and find that older cohorts display better performance than younger ones up to the age of 30, after which the difference is reversed. Controlling for education, young cohorts display consistently better results, demonstrating that education plays an important role.

- Skill mismatch:

    o Measurements of skill mismatch using the diagnostic study of skills, skills usage in the workplace, and self-reported responses on skill mismatches. Pellizzari, M. and A. Fichen (2013)
    o Analysis of skill mismatch, using direct measurement of skills. Desjardins, R. and K. Rubenson (2011).
    o WEF (2014) measures skill mismatches in terms of mismatches by education and skill level, using direct skills assessments and self-reported educational and skill levels required for satisfactory performance in the current job. Some 21% of workers in the OECD countries report over-education, 13% under-education, 33% underutilization of skills, and 13% the need for higher skill levels to do their job.

1.12.   Considerations for users and comments

- PIAAC does not assess socioemotional skills, which limits the range of comparable skills among countries. Some assessments are expected to be included in subsequent rounds, but it is uncertain which skills will be measured or in which round the tests will be administered.
- There are some questions about the low response rates and their impact on the representativeness and validity of the survey results. [Response rate: 1) 75%-70%, four countries: Australia, Ireland, Korea, and the United States; 2) 70%-60% nine countries; 3) 60%-50%, nine countries; 4) 50%-45%, two countries].
- Country participation in this type of initiative is always rewarding, since it is an opportunity to compare results with those of other countries, even in terms of a few skills that are not necessarily high in demand in the productive sector. This type of exercise is also important for countries in the region that are interested in joining the OECD, since it builds bridges of communication between the organization and the countries and generates useful information for country studies that are comparable with those of member countries.
- PIACC is constantly updated and revised; meaning that many of its current deficiencies may be corrected in future rounds.

2. STEP: Skills Toward Employment and Productivity

2.1. Organization responsible for development and/or execution:

World Bank
http://microdata.worldbank.org/index.php/catalog/step/about

2.2. General Objective:

The objective of this initiative is to measure skills in low- and middle-income countries to obtain a better understanding of key issues for policy-making. STEP provides information on skills requirements in the labor market, linkages between skills acquisition and educational achievement, social background, and personality, as well as linkages between skills acquisition and living standards, reductions in inequality and poverty, social inclusion, and economic growth.

Specifically, the survey seeks to determine: 1) the current level and distribution of cognitive, technical, and non-cognitive skills in adults in low- and middle-income countries; 2) the extent of the mismatch between adult skills and employer needs; 3) the relationship between an individual's skills and performance in the labor market; and 4) interventions that countries should consider to boost employment and productivity.

2.3. History

First wave: Data collected between March 2012 and July 2014 in 12 countries: Armenia, Bolivia, Colombia, Georgia, Ghana, Lao PDR, Sri Lanka, Vietnam, Yunnan Province in China, Azerbaijan, Macedonia, and Ukraine

2.4. Target population

Randomly selected population aged 15-64 in randomly selected households in urban areas of low- and middle-income countries.

2.5. Type of skills

Cognitive: Self-reported literacy and numeracy and standardized test of reading proficiency and reading components.
Socioemotional: personality, behavior, and risk preferences. (Short Big Five Inventory - BFI-S), 3 items to assess grit, 2 items to assess hostile attribution bias, and 4 items to assess decision-making)
Job-relevant: Current work-related activities and skills for future employment.

2.6. Content

Household-based survey of individuals and employer survey.

2.6.1. Household-based survey of individuals

A randomly selected individual in the household.

- Background questionnaire:

Collects information on dwelling characteristics and the individual (education, training, experience as apprentices in the formal and informal sector, experience: job transitions, including the first job and information on previous employment, success of the self-employed, family structure, health (BMI and chronic disease, distance to school, etc.).

Self-reporting of cognitive skills: Questions about the use of reading, writing, and numerical skills (mathematical tasks) in the workplace or outside of it. The questions were taken from the January 1991 supplement to the Current Population Survey. It contains eight 8 questions that measure the weekly frequency of reading, writing, math, and computer use, similar to those used in the employer survey in four cities (Holzer, 1996) and in NALS (National Assessment of Adult Literacy) (Sum, 1999), which, in turn, served as the basis for PIAAC. Similar to the STAMP survey, STEP contains questions about the complexity of tasks in these fields.

Self-reporting of job-relevant skills[20]**:** Frequency of problem solving at work, degree of interpersonal contacts in the current job, responsibility for giving formal presentations, degree of independence and supervision; diagnostic question about repetitive work and continuous learning, degree of physical effort, machinery use, inventory of technology use in the workplace, and use of computers outside of work, relevance of educational level in current job, and job-seeking skills.

Self-reporting of socioemotional skills: Includes 31 diagnostic questions on behavior and personality and 7 on risk preferences. The behavior and personality questions consist of psychometric measurements based on the Big Five (3 questions each), Grit (3 questions), "hostile attribution bias" (2

---

[20] There is no skills test. Instead, two complementary selection criteria are used, one of them based on the Classification of Occupations (title, description of general activities), and the other, based on the content of the work activities (self-reported: based on tasks, technology use, and the trilogy of data, people, and things used by the DOT, necessary qualifications for the job, and length of apprenticeship. The advantage of using the International Standard Classification of Occupations is that it is internationally comparable; however, it measures skills associated with a specific job, which do not always match those of workers in those jobs. Another disadvantage is that it does not distinguish between levels of experience in the same job. The job-relevant tasks approach is novel but is increasingly used in the United States (DOT and O*NET), the United Kingdom (UK skill surveys), and Germany (German Qualification and Career Surveys).

questions), and decision-making (4 questions). To measure preferences, simple questions were used to assess attitudes toward risk.

The selection criteria for the questions included were the degree of applicability to populations with low literacy levels (reading comprehension) and little experience with questionnaires, evidence of validity and confidence, evidence of the predictive validity of indicators of education, the labor market, and, especially in extensive surveys, duration.

- Skills measurement module:

Cognitive/reading: Standardized test of reading skills developed by the Educational Testing Services (ETS). It measures reading comprehension and the use of information, using questions and situations from daily life. (The individual must read a prescription label, answer a question about its content, and then list 3 situations in a physician should be consulted). It uses the same scales as PIAAC to ensure comparability between the two surveys.

The test consists of two questionnaires, one of them basic and administered to all participants, and the other consisting of exercises and administered only to individuals with a minimum number of correct answers in the basic instrument. The basic questionnaire contains two sections: 1) Reading components–10 minutes, which measures the reading comprehension process in detail, including imprinted vocabulary, sentence processing, and comprehension of complete passages; 2) A diagnostic literacy test, evaluated by the interviewer, that takes roughly 7 minutes; and 3) an extended test of 18 exercises/questions that cover the same areas assessed by PIAAC (approximately 28 minutes).

2.6.2. Employer survey

The purpose of the employer survey is to determine the structure of the workforce, the skills currently used, the skills that employers look for when recruiting new workers, the willingness of employers to provide training for their workers and the frequency with which they do so, and the association between workforce skills, compensation, and opportunities for promotion.

The sample was taken from employer records, which facilitates the identification of firms and the creation of a representative sample of registered firms. This is less expensive than other sampling methods, such as the use of a business census, a door-to-door search, or the use of the household survey sampling framework. This sampling strategy does not guarantee a representative sample of employers in countries with high levels of informality and, thus, is not representative of the situation in those countries. In each firm, two occupations were randomly selected to ask workers specific questions.

The background questionnaire includes questions about the employer in terms of finances, clients, innovation, skills used by current employees (What is the highest level of computer use in this job?), the

role of skills in decisions, training, compensation, and promotions; the types of skills required when recruiting new workers and the constraints involved in hiring workers with the necessary skills and in training and compensation.

STEP differs from other employer surveys, such as the Investment Climate Survey (ICS), because it concentrates on the type of skills and the perception of their value and contribution to productivity.

The information on skills usage in the workplace is obtained through questions about the regular tasks of workers. The survey measures skills in the same dimensions addressed in the survey of individuals (cognitive skills, behavior, personality, and job-relevant skills). The survey characterizes differences in skills usage between experienced workers and new hires.

2.7. Evaluation

For the tests of socioemotional skills (Big Five, grit, hostile attribution bias, and decision-making) the average response for each scale is calculated, assigning scores using the following scale: 4 for "almost always," 3 for "most of the time," 2 for "some of the time," and 1 for "almost never" (the scale is reversed for negative questions).

These scales showed low levels of confidence (from Cronbach alphas of 0.14 in Vietnam up to 0.47 in Bolivia). According to Angela Duckworth, these low confidence indicators are the result of the combined effect of three factors: The low literacy level of the respondents and their unfamiliarity with self-reporting, the use of reverse-coded items, and the use of four response options rather than the five of the original scales. Notwithstanding, the survey's technical paper mentions that the distributions were not considered problematic and that overall, the predictive validity pattern was considered satisfactory; therefore, the scales were used.

The reading proficiency test developed by ETS rescales the responses using item response theory (IRT). This technique makes it possible, after making some assumptions, to use the same scale to compare individuals with similar levels of knowledge who respond to different questions. This technique also allows for comparisons between surveys, such as those we seek to establish between STEP and PIAAC.

To ensure the comparison between STEP surveys, the parameters estimated in PIAAC 2012 (common IRT item parameters) were used. This guarantees that linkage can be established between the scales of the two surveys and inference structures remain intact.[21]

---

[21] In some countries, it was necessary to implement a single parameter. Once evaluated in each country, population/latent regression models and a series of separate context variables were used separately for each country to produce plausible variables for reading proficiency. See Pierre (2014), pp. 46.

2.8. Basic requirements for administration

The two surveys are administered on paper and require trained personnel to collect the data.

Time: Employer survey of 45-60 minutes and household-based survey of 120-150 minutes (30 for background and 90 for the skills modules; the module for direct measurement of reading skills takes an average of 35 minutes to complete.)

Estimated cost to countries:
Household-based survey: For an average of 3,500 households, the estimated cost is US$470,000-530,000, including implementation (data collection, local consultants) and technical support (supervision of survey implementation and administration of the ETS cognitive module). Implementation takes at least 18 months; about one month is needed for training and translation of the ETS module.

Employer survey: for an average of 400 firms, the estimated cost is approximately US$111,000. Implementation takes about 9-12 months, 4-5 months of which is devoted to data collection.

2.9. Relevance for Latin America

The survey is designed for low- and middle-income countries, and while it does not include employers in the informal sector, which represent a substantial portion of total employers in the countries, it does have modules that are relevant for these countries (for example, household information modules, socioemotional skills assessment modules tailored to populations with high illiteracy rates, etc.).

A key issue is the low confidence of the tests, especially because the low literacy of the population is precisely one of the reasons for it. Worth considering is the possibility of developing socioemotional skills assessments that are reliable in the context of target populations with low education and literacy levels.

2.10.   Administration

Like the other surveys conducted by multilateral organizations, the purpose of STEP is to provide information, but with the goal of using it to influence public policy-making. This survey is of relatively recent vintage (the data was officially published in July 2014), and information about the first experiences with its administration is just coming out.

Bodewig and Badiani-Magnusson (2013) used the STEP results for the annual report on Vietnam, which examines the skills development process in the country and its evolution with respect to the country's high economic and productivity growth since the 1990s. The report stresses the need to develop a

trained workforce using a holistic approach that combines education with acquisition of the new skills needed for modern jobs.

Del Carpio et al. (2013) used the survey information for a technical analysis of the situation in Lao PDR regarding the skills needed to obtain quality jobs. Dundar et al. (2014) conducted a similar analysis for Sri Lanka, while Liang and Chen (2013) concentrated on Yunnan Province in China.

2.11   Considerations for users and comments

- In principle, this survey should be more relevant for developing countries, since it was designed to meet the needs of such countries (operations of labor markets, informal employment levels, skills demand that differs from that of the developed countries, as well as different levels of workforce skills, pronounced differences between rural and urban areas).
- A basic difference between STEP and other international surveys is that it provides information on socioemotional skills. This is relevant, because it has been shown that this type of skill is important for success in the labor market and is highly in demand in the productive sector of countries for which information is available. (See Bassi et al. 2012 for the case of the region.)

The STEP survey is the first attempt to gather this type of information and systematically assess job-relevant skills in developing countries. Items from the STAMP survey were chosen and adapted to ensure their relevance in the context of a developing country. For example, the list of technology tools was restricted to items commonly used in such countries. Moreover, given the possibility that the STEP survey would be implemented in a rural context, job-relevant skills from the agricultural sector were also included.

3. AHELO-Assessment in Higher Education Learning Outcomes

3.1. Organization responsible for development and/or execution

OECD–Skills beyond school

http://www.oecd.org/edu/ahelo

3.2. General Objective

AHELO is part of an initiative aimed at determining the scientific and practical feasibility of assessing the learning and capabilities of university students on graduation.

The objective is to compile information and develop measurements of student learning, forwarding it to institutions of higher education so that they can more accurately pinpoint areas that need improvement. Rather than a ranking, it was developed to provide a diagnosis of worldwide student performance that is valid for different cultures, languages, and types of institution. It was also designed to determine the extent to which institutions of higher education contribute to young people's learning. To this end, an analysis of value added, or learning gains, has been proposed. This component was not included in the feasibility study but was considered in the meetings.[22]

This study yielded good results, but it has not been decided whether the survey will be conducted.

3.3. History

The AHELO project was launched in 2008 as an initiative to assess the quality of higher education. Up to now, the project has concentrated on determining whether it is possible to measure the extent to which institutions of higher education contribute to educational achievement.

AHELO is similar to PIAAC in that it is computer-based and covers a similar subgroup of the population. While PIAAC focuses on the pool of generic skills and outcomes for individuals in the labor market and draws comparisons based on individual and geographical characteristics, AHELO compares institutions in terms of how prepared students are on graduation. It compares programs in each institution and measurements of value added based on variables of interest. Its ultimate objective is to provide information to boost institutional efficiency.

This initial survey was conducted in 17 countries (Abu Dhabi, Australia, Belgium, Colombia, Egypt, Finland, Italy, Japan, Korea, Kuwait, Mexico, the Netherlands, Norway, Russia, the Slovak Republic, Sweden, and the United States) and included around 10 institutions of higher education in each.

3.4. Target population

Youth in their last year of university (approaching the final year of their 3-4-year undergraduate program). A total of 23,000 students and 248 universities voluntarily participated in this feasibility study.

3.5. Type of skills

AHELO measures three types of skills: generic skills and discipline-specific skills in engineering and economics.

---

[22] See Kim and Lalancette (2013) for a review of the literature on measurements of value added in higher education.

The questions are a mix of multiple-choice questions (MCQs) and other constructed response tasks (CRT). The test is administered online.

3.6. Content

The feasibility study measures: generic skills common to all students (critical thinking, analytical reasoning, problem solving, and written communication); discipline-specific skills in engineering and economics; and context, through 3 questionnaires (students, faculty, and institutions).

The generic skills test consists of 2 performance assessment activities (involving writing an essay and answering multiple-choice questions) that require the use of an integrated skill set (critical thinking, analytical reasoning, problem solving, and written communication). Students are asked to answer open-ended questions about hypothetical, but realistic, situations and gather evidence from the information sources provided, including letters, memorandums, research reports, maps, diagrams, tables, etc.

The discipline-specific skills test in economics measures knowledge of the field (key concepts in microeconomics and macroeconomics), applications to real-life problems (using economic reasoning and the application of the discipline's analytical methods to specific issues), effective use of relevant data and quantitative methods (use of quantitative and qualitative data from primary and secondary sources), effective communication and explanations using economic arguments. Finally, it measures independent learning skills, such as thinking reflectively and critically and proposing and conducting research on specific problems.

The engineering test measures generic engineering skills in terms of effective communication and awareness of the broader civil engineering context. It also measures basic engineering concepts such as the scientific and mathematical principles underlying civil engineering (general sciences, materials and construction, structural engineering, geotechnical engineering, hydraulic engineering, and urban and rural planning.), analysis in engineering using analytical methods to identify, formulate, and solve problems; design (understanding and applying design methodologies to meet specified requirements), and finally, engineering practice (practical skills and knowledge required for solving problems, conducting research, designing engineering devices, and other processes. It also addresses non-technical aspects of the discipline, such as professional ethics, responsibilities, and the impact of engineering solutions).

3.7. Evaluation – scoring and interpretation

The scientific feasibility of the survey was measured in two main dimensions, validity and confidence.

Validity was measured in 4 subdimensions: 1) *construct validity,* which measures whether the instrument captures a single dimension of students' latent skills (using factor analysis); 2) *content*

validity, which measures the instrument's ability to capture the content and competencies related to the domain of interest. In order to ensure this validity, a consensus among experts was combined with student opinions; 3) *perceived validity* among the stakeholders involved, which is measured using students' commitment to respond to the survey (interest and time), the self-reported efforts of students, their opinion of the study's relevance; and 4) *concurrent validity,* which measures the degree to which the results vary with those of similar tests taken by the students; this is measured through correlations with other tests and other variables that measure student skills, such as self-reported academic performance. The tests in all these dimensions, with the exception of current validity, yielded satisfactory validity levels.

Confidence measures the stability and consistency of the test results to analyze the replicability of the results under different conditions. Confidence indexes were calculated using the possible-values technique (based on student responses) and the final plausible values technique (using context variables as well).

It was concluded that the study is scientifically feasible and therefore, that instruments can be developed that yield reliable and valid results for different countries, languages, and institutional frameworks.

3.8. Basic requirements for administration

The test requires access to a computer with internet access.
For the generic skills test, students have 120 minutes to complete 2 performance activities that involve writing an essay and answering multiple-choice questions.

3.9. Relevance for Latin America

This initiative is a first attempt at measuring workforce skills in specific disciplines. While it is closely connected with institutions of higher education, it is important to bear in mind when assessing the skills of adults who have accumulated human capital in the labor market.

The initiative starts out with the right questions: How do we know that standardized test results reflect the learning that we want to measure? How do we know that these learning outcomes those that are ultimately needed for success in the labor market? Although the instrument is not perfect, the question format and discussion are interesting and relevant for the countries of the region, which need to define and operationalize learning and skills, not only among students pursuing higher education, but among workers in the labor market to ensure their validity for different programs, institutions, subsystems, and cultures.

3.10.   Administration

Not administered to date.

4. ALL-IALS: Adult Literacy and Life Skills Survey- International Adult Literacy Survey

4.1. Organization responsible for creation/execution:

ALL: OECD
http://nces.ed.gov/surveys/all/

IALS: OECD, Statistics of Canada, and ETS.
https://www.ets.org/literacy/research/surveys

4.2. General Objective:

The main objective of these surveys is to assess literacy, focusing on competency development, and more specifically, to provide an internationally comparable assessment of skills associated with the social and economic characteristics of individuals–in particular, the skills required for their effective participation in the labor market, the political process, and their communities.

4.3. History

There are two rounds of the survey, but the documentation treats them as a single round (2003-2008) with 10 participating countries.
Round 1 (2003) 6 countries: Bermuda, Canada, Italy, Norway, Switzerland and the United States, and also the Mexican state of Nuevo León.
Round 2 (2006-2008) 4 countries: Australia, Hungary, the Netherlands, and New Zealand.

ALL is the successor to the famed IALS, which was based on the methodology, scales, and definitions of literacy used in NALS-1992. IALS was administered in 3 phases (1994, 1996, and 1998) in 22 countries and measured 3 scales: Prose literacy, document literacy, and numeracy (quantitative literacy).
An important feature of ALL-IALS and PIAAC is their comparable components. In the United States, for example, the results of ALL, IALS-1994, PIAAC-2011, and even NALS-1992 can be compared.

The idea behind ALL is to offer an improved version of IALS, with better measurements and quality standards to minimize the sources of variability among surveys and thus increase comparability among countries.

4.4. Target population:

Adults aged 16-65.

4.5. Type of skills

The instrument assesses literacy in three domains: (1) reading of continuous texts (prose); (2) reading of non-continuous texts or schematics (present in forms, graphs, tables, etc.); and (3) the use of numbers (quantitative skills).

Prose-related skills are defined as the ability to understand and use information from texts that reveal their structure and purpose (editorials, news, histories, and fiction). Document-related skills are defined as the ability and knowledge necessary to locate and use information in different formats (job application forms, bus schedules, bills, maps, tables, and graphs) that enable the reader to employ different strategies to access and extract information from them. Finally, numeracy-related skills include completing brief tasks with mathematical content and simple arithmetic operations (simple income and expenditure statement, calculating a tip, completing a merchandise order form, or calculating the interest on a loan).

The instrument also has a pilot program that assesses problem-solving skills (subset of countries).

4.6. Content

Four components: background questionnaire, filter test, literacy skills assessment, and questionnaire on numeracy skills.

Background questionnaire: general information about participants, reading habits, familiarity with information and communication technologies, participation in adult education, and self-reported assessments of literacy and health.

Skills assessment: First, there is a filter test consisting of six core tasks. Respondents who pass this first filter test are administered two more questionnaires, each with approximately 45 questions/tasks, drawn from a pool of items grouped in blocks. The items in the literacy questionnaire are drawn from four blocks of questions (30 minutes for each block). The items on the numeracy questionnaire are drawn from two blocks of questions (30 minutes for each block). This questionnaire takes 60 minutes to complete, with questions randomly selected from each block.

4.7. Assessment

Using the methodology developed for IALS, the skill in each domain is rated on a scale of 0-500 points. Each score denotes the point at which a respondent has an 80% probability of successfully completing a task associated with a similar level of difficulty. Based on the judgment of experts, a 5-level skill scale was constructed for each domain. (See Appendix)

4.8. Basic requirements for administration

Background questionnaire (45 minutes) and assessment (60 minutes).
In some countries, the participants receive financial compensation. (United States US$35).
The questionnaire is administered by trained interviewers. The surveys are administered in households selected from a nationally representative sample.

4.9. Relevance for Latin America

This survey is very similar to PIAAC in terms of the target population and general purpose, but not the skills domains measured. PIAAC captures more cognitive components, including problem solving in technology-rich environments. ALL is exclusively a pencil-and-paper exercise and does not require the use of a computer. Therefore, it does not have the adaptive component of PIAAC. PIAAC is the improved version.

4.10.   Administration

For public policy, mainly reviewing the returns of skills.

Hanushek et al. (2011) "General Education, Vocational Education and Labor Market Outcomes over the Life-cycle" NBER Working Paper No.17504. http://www.nber.org/papers/w17504

IALS is used to compare vocational education with general education in 18 countries. It is argued that the benefits of vocational education in the transition from school to work dissipate over time, due to rapid technological change and the lower adaptability associated with vocational education programs. The trade-off between rapid transition and lower adaptability to change is more apparent in countries with high percentages of apprentices. They use the survey to control for skills differences between countries (and cohorts?).

Leuven et al. (2004), Denny et al. (2004), Hanushek and Zhang (2009), Barone and Vande Werfhorst (2011), McIntash and Vignoles (2001), Green and Riddell 2003 and 2013 [use a follow-up of ALL]

http://www.keepeek.com/Digital-Asset-Management/oecd/social-issues-migration-health/literacy-in-the-information-age_9789264181762-en#page8.

4.11.   Considerations for users and comments

- This survey is very similar to PIAAC in terms of the target population and general purpose but not the domains assessed. PIAAC is the improved version and captures more cognitive components, including problem solving in technology-rich environments. ALL is exclusively a

pencil-and-paper exercise and does not require the use of a computer. Thus, it does not have the adaptive component of PIAAC.

5. LAMP: Literacy Assessment and Monitoring Program

5.1. Organization responsible for development and/or execution

UIS-UNESCO (UNESCO Institute for Statistics)
The primary focus of the program, launched in 2003, is literacy.
http://www.uis.unesco.org/literacy/Pages/lamp-literacy-assessment.aspx

5.2. General Objective

The purpose of LAMP is to furnish robust information to policymakers on adult literacy and numeracy skills, contributing to public debate and the design of adult literacy and education programs that will broaden opportunities for individuals and their families, communities, and countries.

LAMP sheds light on the needs of citizens faced with a complex, constantly changing world full of challenges, where opportunities to participate in different facets of daily life are mediated by written material and thus require sustained development of the skills that guarantee its successful use.

5.3. History

Round 1: 2011. 14 countries. In 04/14 only 4 countries had results: Mongolia, Jordan, Palestine, and Paraguay. Missing were Vietnam, Nigeria, El Salvador, Morocco, Namibia, Afghanistan, Jamaica, Lao PDR, Nigeria, and India.

5.4. Target population

Adults aged 16-64.

5.5. Type of skills

LAMP examines literacy in three domains: 1) reading of continuous texts (prose); (2) reading of non-continuous or schematics (present in forms, charts, documents, etc.); and (3) numeracy (quantitative skills).

Prose-related skills are defined as the ability to read texts arranged in paragraphs, with indentations and headings, that reveal their structure and purpose. Document-related skills are defined as the ability to read non-continuous texts in different formats (tables, schedules, charts, maps, etc.), where readers can use different strategies to enter and extract information. Finally, numeracy skills include performing brief mathematical tasks that require calculating, estimating, and understanding notions of shape, length, volume, currency and other measures.

An attempt is made to ensure that the tasks used for the assessment are relevant for the respondents in terms of their everyday life–that is, that they are related to home and family, health and safety, community and citizenship, consumer economic situations, work-related situations, and leisure and recreation.

5.6. Content

The survey has 4 instruments: 1) background questionnaire, 2) filter test, (3) reading components, and (4) literacy.

1. Background questionnaire: Contains background that provides information on the socioeconomic characteristics of the respondents, including their use of written materials. This is key to analyzing the information generated by the cognitive tools.
2. A filter test to provide an initial estimate of the expected performance of each respondent in order to decide which set of instruments should be used to obtain a more in-depth picture of their skills.
3. Reading components: A module for respondents with lower expected performance that will more accurately indicate their proficiency and explore obstacles to better performance.
4. A module for respondents with higher expected performance that, again, more accurately indicates their proficiency.

5.7. Assessment

The LAMP results show proficiency in the 3 domains at 5 skill levels.

Level 1: Very low. Respondents do not understand the basic information in a text.

Level 2: Low. Respondents can handle only simple materials in which the tasks involved are only moderately complex. Identifies people who can read and have developed coping skills to meet the needs of everyday life but have difficulty meeting new demands (such as learning new job skills).

Level 3: Level of skill formally required for successful secondary school completion and admission to tertiary educational institutions. According to the OECD, this is the minimum level required to meet workplace demands and cope with everyday life in a complex, advanced society.

Levels 4 and 5: Respondents who have demonstrated a command of higher-order information processing skills.

5.8. Basic requirements for administration

The implementation of LAMP is tailored to the characteristics and needs of each country.

The process typically involves the following phases: 1) the exploratory phase (which varies in length); 2) the preparatory phase (a National Planning Report is prepared to establish: i) the study's national objectives; ii) what the information will be used for; iii) the target population; iv) the sample design; v) the characteristics of data collection; vi) the features of data capture and processing operations; vii) the regulations governing confidentiality and quality assurance; and viii) the composition of the national team and division of labor. A costing spreadsheet detailing the cost of implementing LAMP is also prepared); and (3) the implementation phase (18 and 24 months), consisting of a pre-test, field test (6-9 months), and final phase (around 9-12 months; includes adaptation of the cognitive instruments, the development of new items, guidelines, and manuals for data capture, ethnographic elements, coding, and correction and capture of information, as well as the analysis and reporting of results). The final phase is dissemination (3 months).

Each country is responsible for guaranteeing the resources required for implementing LAMP, although the UIS team is available to lend its support in the negotiation of funding opportunities. Implementation typically requires no less than US$250,000. This figure varies with the sample design and other costing elements. UIS does not require countries to cover any overhead, nor does it charge general fees, because its mission is to provide a service to its Member States. Nevertheless, the budget should include some international expenditures for verification of the instruments, observation or training visits, and technical consultants hired for specific tasks in the country.

5.9. Relevance for Latin America

LAMP is particularly interested in ensuring that countries embrace their responsibility to promote literacy. Its purpose is to contribute to the development of national capacities, building on existing experience and promoting it as the best mechanism for supporting national initiatives. This, in turn, means that it encourages countries to improve their institutional environment.

LAMP is more global than IALS, which focuses primarily on OECD countries. LAMP focuses on developing countries with many linguistic families and writing systems.

It is important for boosting the capacity for data collection, training, and institutional work.
It is useful for studying differences in reading comprehension and reading components in different languages. (The countries of the region have indigenous populations. This survey can be used to guide efforts aimed at understanding the effects of the differences in the language of both respondents and translators.)

5.10.   Administration

Construction of a literate environment index that presents individual and household characteristics. Analysis of reading skills. Description of reading components and linguistic differences (for example, a question is valid if tasks such as spelling, visual word recognition, and oral vocabulary predict literacy). The model for predicting literacy may be common to all languages.

5.11.   Considerations for users and comments

- LAMP is not radically different from other, more complete and recent surveys such as PIAAC and STEP. Therefore, either of these two is preferable if the goal is to measure skills for work and life.
- Unlike PIAAC and IALS, LAMP is more global in nature, emphasizing developing countries with many linguistic families and writing systems. Thus, it is more tailored to the needs of countries with populations characterized by low levels of education, since it is simpler in terms of content and administration.
- An important consideration is the fact that LAMP has been designed to help strengthen national capacities, building on existing experience and promoting it as the best mechanism for supporting national initiatives. This, in turn, means that it encourages countries to improve their institutional environment (important for strengthening the capacity for data collection, training, and institutional work.)
- It is useful for studying differences in reading comprehension and reading components in different languages. This is especially important for the countries in the region that have a high proportion of indigenous populations and differences in dialect/language/idioms.

6. Disconnected: Skills and Trajectory Survey (STS) in Chile and Argentina and Skills Demand Survey in Chile, Argentina, and Brazil

6.1. Organization responsible for development and/or execution:

Part of the Disconnected Project of the IDB-Education Division (Marina Bassi, Matías Busso, Sergio Urzúa, and Jaime Vargas).

6.2. General Objective

To provide information about the current mismatch between the skills of recent graduates of the educational system and the skills needed by the productive sector. Specifically, it looks at the

association between the population's educational level and its performance in the labor market, considering the role that skills play in this process.

As part of this project, surveys of employer demand for skills in the automotive, food, hospitality, financial, and retail sector were conducted in Chile, Argentina, and Brazil.

6.3. History

Developed to bridge the information gap with respect to skill shortages in the region and skills assessment. As of this time, there are no plans to repeat it for other countries or periods.

The Skills and Trajectory Survey was conducted in Argentina between May and June 2010 and in Chile, between October and December 2008. The Skills Demand Survey was conducted in Argentina, Chile, and Brazil between October 2009 and January 2010.

6.4. Target population

The target population of the Skills and Trajectory Survey in Argentina consisted of individuals aged 25-30 in April 2010, and in Chile, urban residents at the time of the survey (October to December 2008)[23].

The target population of the Skills Demand Survey was employers in the automotive, food, hospitality, financial, and retail sectors in Chile, Argentina, and Brazil. The unit of analysis was occupation-employer; thus, in every interview, the respondent was asked to report on three occupations, including the occupation considered most closely related to the company's business.[24]

A total of 1,176 employers were interviewed. The sampling framework consisted of lists of employers used by canvassing companies (telephone directories, industry directories, lists compiled through online searches), since the national statistics institutes did not have an employer census. The goal was to obtain a minimum of 60 employers per sector per country.

6.5. Type of skills

The survey has four scales that measure general intellectual ability and three types of socioemotional skills: social skills, self-efficacy, and metacognitive strategies.

---

[23] In Argentina, the Skills and Trajectory Survey was conducted in three metropolitan areas: Metropolitan Buenos Aires, Córdoba, and Mendoza, and in Chile, in the country's major cities, with urban coverage in every region in the country except Region IX.

[24] Information was compiled on 3,015 occupations.

The three subtypes were selected for their importance to educational outcomes and their influence on performance in the labor market. The feasibility of administering these tests in the context of a household survey (time, availability, etc.) was also considered.

6.6. Content

6.6.1. Skills and Trajectory Survey

The survey includes a background questionnaire on the individual's family and education, together with retrospective questions about experiences in the labor market.

As for skills, general intellectual ability is measured using a test developed by the Department of Psychology of the Pontificia Universidad Católica de Chile. The test consists of a series of eight figural analogies. Figural analogies are typically used to measure fluid intelligence–the type of intelligence closest to general intelligence. This short test showed a high degree of correlation with other, more exhaustive tests of general intelligence.

Scores were calculated on an eight-point scale, where 0 indicated no correct analogy, and 8, all correct. According to the university's Centro MIDE, this test is highly correlated with other measurements of general cognitive ability.

For socioemotional skills, there were tests for each of the selected subtopics. In all cases, the test consisted of a series of questions for which the respondent had to select one of four graduated responses (1. very little, 2. somewhat, 3. Moderately, and 4. strongly). The average of the responses to each test was used as a scale.

Self-efficacy is measured by 12 questions designed to determine an individual's ability to get organized and meet his goals. **Social skills** are measured by a 12-question test that characterizes how individuals relate to others–for example, in terms of communication, leadership, etc. Finally, **metacognitive strategies** (the ability to plan activities) are measured with a set of 12 questions aimed at characterizing how individuals approach various tasks and use critical thinking.

6.6.2. Skills Demand Survey

The Skills Demand Survey contains questions about employer demand for skills, wages, general characteristics of the company, the hiring process, and employer perceptions of skill gaps. The skills considered were:

- Knowledge skills: language and communication (ability to listen, ask questions, and express concepts and ideas effectively); reading and writing skills; the ability to solve basic math problems; and the ability to apply critical thinking to understand and solve problems.

- Socioemotional skills: attitudinal skills (the ability to work cooperatively with others while controlling one's emotions and avoiding negativity); responsibility and commitment to the organization's objectives and to the performance of assigned tasks; and customer service skills (good personal presentation, respect, friendliness, etc.).
- Specific skills: employees' willingness to acquire new skills specific to their job; use of specific machinery and computers.

6.7. Evaluation – scoring and interpretation

The evaluation of the questions in the Skills and Trajectory Survey is described in the previous point.

6.8. Basic requirements for administration

Skills and Trajectory Survey: consists of a 60-minute interview of individuals.

The Skills Demand Survey combined online and personal surveys (by telephone and in person). The survey took approximately 45 minutes.

6.9. Relevance for Latin America

Relevance for the region. Used by other countries as a reference for skills measurement.
It is a required reference on skills shortages in the region and an example of questionnaires for cognitive and socioemotional skills measurement.

6.10. Administration

The survey was created for the book *Disconnected.* The skills module was subsequently implemented, with some variations, for several studies in Colombia and Costa Rica.

6.11. Considerations for users and comments

- This survey was developed in response to information gaps about skills shortages and skills assessment in the region. As of this time, there are no plans to repeat it for other countries or periods.
- It is important to note that great pains were taken in this survey to select skills that can easily be related to the type of skills required by the productive sector, building a conceptual bridge between the skills acquired in school and the perception of what is needed in the workplace.

- Unlike the international studies conducted by such organizations as the OECD, the World Bank, UNESCO, et al., the skills assessment modules are available for uses other than initially conceived for the survey.

Part 2: Batteries of questions, questionnaires, and individual scales

In this section, the reader should refer to the attached table, which contains an extensive list of individual tests offered by commercial assessment companies.

7. IPIP- International Personality Item Pool

7.1. Organization responsible for development and/or execution

International Personality Item Pool: A Scientific Collaboratory for the Development of Advanced Measures of Personality Traits and Other Individual Differences (http://ipip.ori.org/).

This initiative is based on the work of Lewis R. Goldberg, University of Oregon and Oregon Research Institute. (Goldberg, 1999 and Goldberg et al., 2006).

7.2. General Objective

The main objective of this initiative is to create a series of public domain personality inventories that enable users and academics to collaborate for the continuous improvement of the inventories. Access to broad-bandwidth personality inventories is very limited, in contrast to limited-bandwidth scales that, while available to the general public, assess only one specific personality facet or trait.[25] The purpose of this website, therefore, is to breach the access barriers to major personality tests to improve and refine them.

---

[25] Some examples of limited bandwidth scales are: Motivation, conservatism, dogmatism, empathy, extraversion-introversion, guilt, hostility, neuroticism, openness to experience, optimism, self-esteem, trust in others, etc. Some of the most cited examples of broad-bandwidth inventories son: the Minnesota Multiphasic Personality Inventory (MMPI, the California Psychological Inventory (CPI), the Sixteen Personality Factor Questionnaire (16PF), etc. See Table A1 in the appendix for a definition and description of the tests.

7.3. History

This project involves the development of scales similar to the ones most commonly used but require no user fee and promote collaboration to improve the scales over time. The program was developed by Lewis R. Goldberg, of the University of Oregon, and financed by a grant from the National Institute of Mental Health, U.S. Public Health Service (Grant MH-49227).

In the first phase of the project, scales comparable to the main broad-bandwidth scales were constructed: 45 bipolar facets that attempt to reproduce the facets of AB5C26, 30 scales comparable to the NEO-PI-R, 16 scales comparable to the 16PF (Conn and Rieke, 1994), 30 items comparable to the 31 of Cloninger's Temperament and Character Inventory (TCI), and 33 scales comparable to the CPI (Gough, 1996).

The first validity study was conducted by measuring the predictive power of the scales in terms of three behaviors linked with health and behavior: (1) risk prevention; (2) good health practices; and (3) general health concerns. These validity studies confirmed that the IPIP scales had greater predictive power than the original scales.

7.4. Target population

Adults and, for some, youth. The target population varies with the type of battery.

7.5. Type of skills

The IPIP scales and inventories cover non-cognitive and personality skills but include some vocational skills as well. The types of skills assessed are summarized below, classifying the scales into three groups: a) structural scales similar to the Big Five, b) broad-bandwidth personality inventories; and 3) other batteries.

a) Structural scales similar to the Big Five: 5 scales summarized in

b) **Table 8**8 below.

Table 8. Batteries similar to the Big-Five

Scales and Tests[27]	Measures
Broad Big-Five scale (10- and 20-question scales)	Openness
	Extraversion
	Conscientiousness
	Agreeableness
	Neuroticism
7-factor scale -Saucier, G., 1997 (10-question scale)	Big-Five
	Attractiveness
	Negative Valence
Scale comparable to AB5C facets (10- and 20-question scales)	Measures 45 Big Five combinations http://ipip.ori.org/newAB5CTable.htm
Scale of 10 aspects of the Big Five (10-question scale)	Openness: Intellect and openness
	Extraversion or Sociability: Enthusiasm and Assertiveness
	Conscientiousness: Dutifulness and Order
	Agreeableness: Sympathy and Cheerfulness
	Neuroticism: Impulsiveness and Immoderation
8 IPIP-IPC scales (32-question scale)	Assured-Dominant
	Arrogant-Calculating
	Cold-Hearted
	Aloof-Introverted
	Unassured-Submissive
	Unassuming-Ingenuous
	Warm-Agreeable
	Gregarious-Extroverted

Source: http://ipip.ori.org/newMultipleconstructs.htm

c) Broad-bandwidth personality inventories: 14 comparable questionnaires listed in

---

[27] See Table A1 in the appendix for a definition and description of these tests.

d) **Table 9**9 below, which provides the title of the questionnaire, the inventory to which it is comparable, and the type of skills measured.

Table 9. Questionnaires comparable to broad-bandwidth personality inventories

Questionnaire[28]	What it assesses
NEO-PI-R domains (10 and 20 questions)	Openness
	Extraversion
	Conscientiousness
	Agreeableness
	Neuroticism)
NEO-PI-R and 30 NEO-PI-R facets (10, 20, 120 questions)	30 facets derived from the Big-Five http://ipip.ori.org/newNEO_FacetsTable.htm
Questionnaire comparable to 16PF (160 questions)	16 personality traits: Warmth, Intellect, Emotional Stability, Assertiveness, Gregariousness, Dutifulness, Friendliness, Sensitivity, Distrust, Imagination, Reserve, Anxiety, Complexity, Introversion, Orderliness, and Emotionality.
	Big Five traits: extraversion, independence, self-discipline, anxiety, and openness
Questionnaire comparable to CPI (330 questions and 33 scales)	Popular Scales: Assertiveness, Complexity, Sociability, Adventurousness, Insight, Self-efficacy, Depth, Responsibility, Stability, Self-control, Temperance, Tolerance, Optimism, Planfulness, Intellect, Comprehension, Competence, Disorder, Timidity/Forcefulness.
	Structural Scales: Orientation, Self-discipline, Self-efficacy
	Happiness, Calmness, Politeness
	Special indexes: Happiness, Tough-mindedness, Creative Temperament
Questionnaire comparable to HPI (70 questions higher level constructs, 60 questions for occupational scales)	7 Higher-level constructs: Stability, Leadership, Sociability, Friendliness, Dutifulness, Creativity, Quickness
	6 Occupational Scales: Calmness, Happiness, Cooperation, Toughness (meticulousness), Gregariousness, Competence
	44 Homogeneous Item Clusters (HICs): 44 traits/adjectives related to personality.
Questionnaire comparable to TCI (289 questions)	7 dimensions of personality comprised of 4 temperaments (Novelty-seeking, Harm avoidance, Reward dependence, Persistence) and 3 characters (Self-directedness, Cooperativeness, Self-transcendence)
Questionnaire comparable to MPQ (210 true or false questions)	Four major characteristics: Positive Emotional Temperament, Negative Emotional Temperament, Constraint, Absorption.
	11 scales of primary traits: Joyfulness, Power-seeking, Achievement-seeking, Friendliness, Emotional Instability, Belligerence, Distrust, Planfulness, Risk-avoidance, Conservatism, Imagination.
	3 validity scales: unlikely virtues, inconsistencies with true answers and inconsistencies due to variable responses
Questionnaire comparable to JPI-R	5 cluster scores and 15 scale scores
	Analytical: Intellectual complexity, Intellectual breadth, Ingenuity,

[28] See Table A1 in the appendix for a definition and description of the tests mentioned.

(210 true or false questions)	Tolerance
	Extroverted: Empathy, Anxiety, Conformity
	Emotional: Sociability, Social confidence, Activity level
	Opportunistic: Machiavellianism, Risk-taking
	Dependent: Organization, Traditionalism, Responsibility
Questionnaire comparable to 6FPQ (60 questions and 172 questions)	6: Big five plus methodicalness
	18 constructs: Gregariousness, Leadership, Exhibitionism, Docility, Calmness, Adaptability, Conservatism, Deliberateness, Orderliness, Reclusiveness, Unpretentiousness, Self-sufficiency, Adventurousness, Comprehension, Culture, Striving, Resourcefulness, Playfulness.
Questionnaire comparable to HEXACO-PI (60 questions and 172 questions)	24 constructs: Sincerity, Fairness, Greed avoidance, Modesty, Fearfulness, Anxiety, Dependence, Sentimentality, Expressiveness, Social boldness, Sociability, Liveliness, Forgiveness, Gentleness, Flexibility, Patience, Organization, Diligence, Perfectionism, Prudence, Aesthetic appreciation, Inquisitiveness, Creativity, Unconventionality.
Questionnaire comparable to VIA (240 questions)	24 scales: Appreciation of beauty, Capacity for love, Citizenship, Curiosity, Equity, Forgiveness, Gratitude, Hope, Humor, Industry, Integrity, Judgment, Kindness, Leadership, Love of learning, Modesty, Originality, Perspective, Prudence, Self-regulation, Social/personal intelligence, Spirituality, Valor/Bravery, Zest.
Questionnaire comparable to Yarkoni's 181-item AMBI inventory (200 questions)	Measures more than 200 personality traits with 181 questions. Yarkoni (2010) proposes this test, which captures the personality dimensions that are measurable using broad-bandwidth inventories (NEO-PI, HEXACO, TCI, HPI, 6FPQ, CPI, JPI-R, and MPQ) http://ipip.ori.org/AMBIComparisonTable.htm

e) Other batteries

- Simms' Personality Disorder Scales
- Potential "Emotional Intelligence" Components (Barchard, 2001)
- Constructs in Gray's Behavioral Inhibition and Activation Systems (BIS/BAS: Carver and White, 1994)
- Oregon Vocational Interest Scales (ORVIS) (Pozzebon, Visser, Ashton, Lee, and Goldberg, 2010)

7.6. Content

The website contains an extensive list of batteries, tables for comparison with other standardized tests, and keys to scoring and evaluation. The most commonly used instruments are listed below, but it is important to refer to the website, because it is constantly being updated.

a) Five structural batteries similar to those of the Big Five: Big Five in 10 domains, scale of 7 constructs, 45 facets of AB5C, 10 Big-Five aspects, CPI scales for 8 IPIP (See

b) **Table 8**8)
c) 14 broad-bandwidth personality inventories: NEO-PI-R domains and facets, Johnson's 120-item IPIP NEO-PI-R, 16PF, CPI, HPI, TCI, MPQ, JPI-R, 6FPQ, HEXACO-PI, VIA, Yarkoni's 181-item AMBI inventory (See Table 9)
d) Other batteries (See previous section).
e) Four single-construct tests:

1) Personal Attribute Survey (PAS): Self-esteem (Rosenberg, 1965), self-monitoring (Snyder, 1974), self-consciousness in public and private (Buss, 1980), optimism and disposition (Scheier, Carver, and Bridges, 1994), impression management (Paulhus, 1991), self-deception (Paulhus, 1991), internal locus of control (Levenson, 1981); self-perceived attractiveness.

2) Comprehensive Health Survey (CHS): Need for Cognition (Cacioppo and Petty, 1982), Cognitive Failures Questionnaire (Broadbent et al., 1982), Obsessive-Compulsive Inventory (Foa et al., 1998).

3) Personal Reactions Survey (PRS): Attention and hyperactivity (Span et al., 2002), Obsessive-Compulsive Inventory (Foa et al., 2002), hypomanic characteristics (Eckblad and Chapman, 1986), Sensation-seeking facets (Hoyle et al., 2002).

4) Personality, Emotions, and Attitudes Survey (PEA): measures alexithymia (Taylor et al., 1986), borderline personality disorder (Leichsenring, 1999), fantasy proneness (Merckelbach et al., 2001), hypochondria (Katz and Zenger, 1999), magical ideation (Eckblad and Chapman, 1983) psychopathies (Levenson et al., 1995): 2 subscales, somatoform dissociation (Nijenhuis et al., 1997).

7.7. Evaluation – scoring and interpretation

Provides sample questionnaires and all related questions. See http://ipip.ori.org

7.8. Basic requirements for administration

Few requirements, easy to administer and score.

7.9. Relevance for Latin America
Free, easy to take and administer, although it may require technical personnel familiar with basic statistics to do the calculations.
The authors do not know of specific studies that have used these tests for research on interventions or programs in the region, but STEP examined the program in depth to include it in its initiative.

7.10.   Administration

There have been few instances of the administration of these instruments in the region to analyze aspects of the labor market. However, the technical committee of the World Bank's STEP program carefully reviewed these tests for the survey's preliminary study.

In addition, there are Spanish translations that have been validated in several countries in the region.

- IPIP-VIA: A Spanish translation with confidence and validity tests is available. Used to explore the relationship between the IPIP-VIA scales and assessments of psychosocial well-being. Contact: Merche Ovejero, European Institute of Positive Psychology, Madrid (mercheovejero@iepp.es)
- IPIP-16PF: Psychometric studies from the Spanish version of the battery are available. Contact: Prof. Edgardo Pérez, Facultad de Psicología, Universidad Nacional de Córdoba, Argentina (edrape@onenet.com.ar). http://www.revistaevaluar.com.ar, No. 4.
- IPIP-16PF: Uses the Spanish version for the selection of personnel, negotiation, and conflict resolution. Contact: Prof. Richard A. Posthuma, College of Business Administration, University of Texas, El Paso (rposthuma@utep.edu)
- Big Five: Translation of the 50-question battery, validated in Mexico. Used to analyze the influence of personality traits on customer satisfaction with the mobile phone service. Contact: Rodrigo de Oliveira, Telefónica Research, Barcelona, Spain. (oliveirard@yahoo.com.br).

The complete list of publications that use the batteries developed by this initiative can be found at the following links:

http://ipip.ori.org/newPublications.htm

http://projects.ori.org/lrg/

8. NIH Toolbox

8.1. Organization responsible for development and/or execution

National Institutes of Health, United States

http://www.nihtoolbox.org

8.2. General Objective

This battery was created to overcome the main barriers of the other batteries available on the market, increasing their accessibility and facilitating comparisons between studies. There are countless tests and batteries of questions on the market, but they are generally expensive, lengthy, and must be

administered by highly specialized personnel. The most commonly used batteries, moreover, are not uniform, making comparisons between studies and disciplines even harder.

The Toolbox is an integrated series of tests that measure four skills dimensions: cognition, sensation, emotional, and motor. The tests have been validated for use with different cultures, ethnicities, geographic groups, ages (3-85 years), and types of studies. The Toolbox is inexpensive, requires no payment of patent royalties, and uses state-of-the-art psychometric methodologies and computer-adapted tests that can be calibrated to the respondents' skill level. The entire test takes 2 hours and is available in English and Spanish.

8.3. History

The subdomains included in the measurement of each skills domain were selected on the basis of extensive literature reviews, interviews with experts, and many formal requests for information by the NIH-funded researchers. Thus, after lengthy sessions for consensus building, meetings, and seminars, agreement was reached on the tests and subdomains to be included in the final version of an initial series of more than 1,400 measurements for potential inclusion. The selection criteria included their applicability throughout life, validity and psychometric soundness, brevity, ease of use and administration, applicability to different settings and different groups, and intellectual property barriers. There was also a preference for instruments that had already been validated for use with individuals aged 3-85.

8.4. Target population

The batteries are designed for different age groups ranging from 3 to 85 years of age, depending on the skills dimension to be assessed. The tests have been validated for use with different cultures, ethnicities, and geographic and age groups.

Cognition and motor skills: 7 years and over. The battery includes a special module for children aged 3-6. Emotional skills: designed for individuals aged 8 and over. A module for the parents of children aged 3-12 is also included. The majority of the emotional health instruments differ by age group (generally 8-12, 13-17, and 18+). The sensation battery is recommended for individuals aged 7 and over, although some instruments are available only for certain age ranges (for example, batteries that assess taste are for individuals aged 12 and over, and those for pain, for individuals over 18). Specialized scales are also available for different age ranges.

8.5. Type of skills

This is an integrated set of batteries that assess four skills dimensions: cognition, sensation, emotion, and motor. A broader definition of each dimension is presented below.

- Cognition

Cognition assessments are done to measure activities that contribute to brain health. Although this objective appears to be far removed from the assessment of workforce skills and competencies, it includes assessments of a number of skills that are indispensable in the workplace, such as language (which in many other surveys is the only measurement of cognitive ability), activities planning and execution skills in terms of meeting objectives, and various aspects of memory and attention.

These assessments measure the mental processes needed for knowledge acquisition and comprehension, such as thinking, understanding, remembering, judging, and solving problems. These advanced brain functions include language, imagination, perception, and the planning and execution of complex behaviors.

This battery measures six subdimensions: executive function, attention, episodic memory, language, processing speed, and working memory. Administration of this battery yields the following scores: cognitive function composite score, fluid cognition composite score, and crystallized cognition composite score.

There is also a battery of questions to assess cognitive skills at early ages (3-6 years) that yields an early childhood cognitive function composite score.

- Emotion

This group of instruments was created to emphasize the importance feelings, thoughts, and behaviors to an individual's overall health and well-being. It has also been found that several of the aspects they measure are important for job performance, especially those related to stress, self-efficacy, and negativity.

Emotion is an affective state in which feelings such as happiness, sadness, or fear are experienced by individuals, making it different from other states such as cognition. This battery has four main subdomains: psychological well-being, social relations, stress and self-efficacy, and negativity. The assessments include a version for self-reporting and, for certain ages, versions for parental reporting.

- Motor

The purpose of the measurements developed to assess motor skills is to determine the overall physical health of individuals, their ability to perform activities of daily living, and their general health. Thus, they measure an individual's muscle strength, ability to manipulate small objects quickly, balance, typical gait

speed, and endurance. These skills also serve as a general health assessment and are useful for assessing skills in different types of occupations.

The motor skills batteries are designed for individuals aged 7 and over, although they include a special battery to assess early childhood development. Motor skills include five subdomains: dexterity, grip strength, standing balance, gait speed, and endurance.

- Sensation

The instrument measures different dimensions of sensory skills: audition, visual acuity, vestibular balance, olfaction, taste, and pain.

8.6. Content

8.6.1. Cognition

This battery measures six subdomains, described below. Each subdomain has at least one related instrument or questionnaire. There are nine instruments in all, only seven of which are used to calculate the cognitive function composite score (the instruments excluded from the calculation are the NIH Toolbox Auditory Verbal Learning Test (Rey) and the NIH Toolbox Oral Symbol Digit Test).

Executive function: Measured by two instruments: 1) NIH-Toolbox Flanker Inhibitory Control and Attention Tests, and 2) NIH-Toolbox Dimensional Change Card Sort Test-DCCS.

Attention: Measured by a single instrument: NIH Toolbox Flanker Inhibitory Control and Attention Tests.

Episodic memory: Measured by two instruments: 1) NIH Toolbox Picture Sequence Memory Test and 2) NIH Toolbox Auditory Verbal Learning Test (Rey)

Language: Measured by two instruments: 1) NIH Toolbox Picture Vocabulary Test and 2) NIH Toolbox Oral Reading Recognition Test".

Processing speed: Measured by two instruments: 1) NIH Toolbox Pattern Comparison Processing Speed Test and 2) NIH Toolbox Oral Symbol Digit Test.

Working memory: Measured by the NIH Toolbox List Sorting Working Memory Test.

8.6.2. Emotion

This battery measures four subdimensions: psychological well-being, social relationships, stress and

self-efficacy, and negative affect. The tests are brief pencil-and-paper questionnaires (in some cases adapted for computer use) for measuring skills through self-reporting and subsequent diagnosis or parental reporting (depending on the age of the respondent). The majority of emotional health instruments differ with the age range (generally 8-12, 13-17, and 18+).

Psychological well-being: Measures three components 1) Positive affect, through the NIH Toolbox Positive Affect Survey; 2) Life satisfaction, through the NIH Toolbox General Life Satisfaction Survey; and 3) the NIH Toolbox Meaning and Purpose Survey.

Social relationships: Concentrates primarily on four aspects of social relationships: 1) Perceived social support, measured by the NIH Toolbox Emotional Support Survey and the NIH Toolbox Instrumental Support Survey; 2) Companionship, measured by four surveys: NIH Toolbox Friendship Survey, NIH Toolbox Loneliness Survey, NIH Toolbox Social Withdrawal Survey, and NIH Toolbox Positive Peer Interaction Survey; 3) Social distress, measured by three surveys, NIH Toolbox Perceived Rejection Survey, NIH Toolbox Perceived Hostility Survey, NIH Toolbox Peer Rejection Survey, and 4) Positive social development, measured by one survey, the NIH Toolbox Empathic Behaviors Survey.

Stress and self-efficacy: Measures two areas related to stress and adaptive capacity–the former through the NIH Toolbox Perceived Stress Survey and the latter through the NIH Toolbox Self-Efficacy Survey.

Negative affect: This subdimension has three components. The first is anger, measured by four questionnaires: NIH Toolbox Anger-Affect Survey, NIH Toolbox Anger-Hostility Survey, NIH Toolbox Anger-Physical Aggression Survey, and finally, NIH Toolbox Anger Survey. The second is fear, for which five instruments are used: NIH Toolbox Fear-Affect Survey, NIH Toolbox Fear-Somatic Arousal Survey, NIH Toolbox Fear-Over Anxious Survey, NIH Toolbox Fear Survey, and NIH Toolbox Separation Anxiety Survey. The third and final component is sadness, measured by the NIH Toolbox Sadness Survey.

All of these instruments have supplemental measures tailored to different age ranges and have different formats, especially when the format of the original instrument is adapted and requires the use of a computer.

8.6.3. Motor

The motor skills batteries are designed for individuals aged 7 and over, although they include a special battery to assess early childhood development. Motor skills consist of five subdomains: dexterity, grip strength, standing balance, gait speed, and endurance. Manual dexterity is measured by the NIH Toolbox 9-Hole Pegboard Dexterity Test. Strength is measured by grip strength with the NIH Toolbox Grip Strength Test. Balance is measured by the NIH Toolbox Standing Balance Test. Gait speed is measured by the NIH Toolbox 4-Meter Walk Gait Speed Test, and finally, endurance is measured by the NIH Toolbox 2-Minute Walk Endurance Test".

8.6.4. Sensation

Sensation tests use different instruments to assess audition, vision, vestibular capacity, olfaction, taste, and pain.

8.7. Evaluation – scoring and interpretation

Scores are calculated using classical test theory[29], which consists of calculating the average of the normalized scores for each of the measurements made to assess the same aspect.

There are also scores that are age-adjusted, fully adjusted, not adjusted (median =100 and standard deviation=15) and national rankings in percentiles (calculated using the age-adjusted scores).

Administration of the cognitive skills battery yields three scores: cognitive function composite score, fluid cognition composite score (all tests except the two language tests), and crystallized cognition composite score (includes only language tests: picture vocabulary and oral reading recognition).

Emotional skills test scores are calculated for each of the measurements and no composite score is provided.

8.8. Basic requirements for administration

Instructional materials are available for the administration of this battery. They consist of a 98-minute video, a brief manual, a detailed manual with instructions for administering the tests, and other related resources.

The instructional video is available free of charge at: http://www.nihtoolbox.org/WhatAndWhy/Assessments/E-learning%20files/player.html

---

[29] Individuals take a test, and the final score is used for comparison purposes. A higher score is associated with a higher level of the skill assessed. Under this theory, the test measures the skill plus an error component that depends on both the individual's innate skill and the characteristics of the test, components that cannot be separated.

Another resource is www.assessmentcenter.net, which offers a free online tool for accessing not only the tests developed by NIH (NIH Toolbox) but also an extensive list of tests through the Patient-Reported Outcomes Measurement Information System (PROMIS).

Time

The emotional skills tests take an average of 12 to 22 minutes, and the length of time for the other tests varies, but the entire test is guaranteed not to exceed 2 hours.

Costs

The costs of this test derive basically from the costs associated with technical and user support. This includes access to an online help tool, assistance with the use of the online tools and formal implementation of the studies, and technical support for all the tests included. The costs ranges from US$1,500 per year to assess 100 patients or less and US$5,000 for studies that exceed this threshold.

These costs do not include the costs associated with the equipment and software needed for the tests, which include internet access, computers, basic software, and in some cases, a large monitor. A detailed description of these items can be found at:

http://www.nihtoolbox.org/WhatAndWhy/Materials/Pages/default.aspx

8.9. Relevance for Latin America

This instrument is relevant for the region, as it has been developed with the highest standards of quality. To date, it has not been administered in the region. Some materials–the technical manual, for example– are available in Spanish, and although the tests were validated with a representative group of respondents in the United States, the group included Spanish-speakers. Language may be a barrier to its administration, because several of the instructional manuals and videos are only in English.

8.10. Administration

Despite the long list of publications, the Toolbox has not been used in economic or labor market contexts. Rather than a disadvantage, this is an opportunity for initiatives in the region.

http://www.nihtoolbox.org/Publications/Pages/Articles1.aspx

9. CPS-Social and Personal Competencies Scale

9.1. Organization responsible for development and/or execution

Inter-American Development Bank based on the work of Mayra Brea

9.2. General Objective

This is a test of socioemotional skills developed to objectively assess the impact of the DBC (development of basic skills) Module on young people who received assistance from the Youth and Employment Program in the Dominican Republic.

The objectives of the DCB are to develop positive values, attitudes, and basic skills (basic competencies included in the Training Guide) among young people living in poverty, improving their chances of finding a good job and mitigating the social risks imposed by the educational exclusion and marginalization that they experienced early on.

9.3. History

The CPS scale, created in July 2010, originally had 70 questions and 11 dimensions. It was based on the academic and international literature on human competencies and drew on successful experiences with the DCB Module of the Youth and Employment Program in the Dominican Republic, implemented with the Training Guide for Basic Skills Development. The current revised and standardized CPS (June 2011 Report) has 44 questions.

9.4. Target population

Young adults aged 16-29.

The CPS scale has been field tested, adapted, and validated in its construct in the Dominican Republic (See first, second, third, and fourth report of Mayra Brea, World Bank consultant, July-August 2010) in a socially marginalized, low-income population of young adults, both male and female, who participated or were connected in some way to the Youth and Employment Program. The language and phrasing of the items had been tailored to this population.

9.5. Type of skills

The current revised and standardized CPS assesses six basic skills: leadership, behavior in situations of conflict, self-esteem, ability to relate to others, organization-order, and empathy and communication skills. A description of these basic skills, as described in Brea, 2011, is presented below.

Leadership: The ability to influence peers and obtain their commitment to a common goal. It implies the ability of young adults to impact others and make themselves known to or admired by their peers. A leader is always willing to participate actively in matters important to his community, has the ability to work with others and commit to the goal of a team, and can harmonize agreements and activities with others.

Behavior in situations of conflict: The ability to recognize, express, and properly control emotions and to think before acting, along with the ability to recognize the source of an interpersonal or social conflict, consider the points of view involved, and propose solutions or remedies to any problems and situations

that arise. It entails the ability to weigh potential alternatives and verify the relevance of the solutions, as well as a certain degree of responsibility toward obligations.

Self-esteem: Recognition of one's own abilities, personal satisfaction. A positive or negative self-image.

Ability to relate to others. The ability to create and maintain social connections with peers and others in one's environment. It entails knowing how to behave in a social situation in order to handle oneself with self-assurance in a manner consistent with one's goals and points of view.

Organization-Order: The ability to plan ahead and the willingness to maintain order among the tools and materials used in daily living. It also involves a certain commitment to the goals set by the work team and the social setting.

Empathy-Communication skills: The ability to understand and accept others, to put oneself in their shoes, to be open to other points of view, and to be respectful of people, ideas, values, and/or customs that differ from one's own. Also, the ability to accurately and reliably express and understand ideas or messages, which can contribute to good relations and fitting in.

9.6. Content

	CPS Dimensions	Questions												# items
1	*Leadership*	12	25	27	29	38	39	45						7
2	*Behavior in situations of conflict*	5	6	8	15	21	24	26	30	41	42			10
3	*Self-esteem*	4	13	32	36									4
4	*Ability to relate to others*	17	20	22	31	37								5
5	*Organization-Order*	7	11	14	19	33	35							6
6	*Empathy-Communication skills*	3	9	10	16	18	23	28	34	40	44	46	47	12

Source: Brea, 2011.

9.7. Evaluation – scoring and interpretation

This battery consists of 44 multiple-choice questions with four response options: "Nothing like me", "Not much like me", "Like me", and "Very like me". Each of these responses is assigned a score of 0 to 3. Negative affirmations take inverse values.

The score is calculated using the sum of the items in each dimension. The score obtained is compared with the Standard Normal Table, where the corresponding range, or Sten, is observed and thus, its significance for interpreting the data.

Stens are values on a normalized scale of 10 standard points, which have the following meanings: 10: Very high; 9: High; 8: Normal-High; 4 to 7: Normal; 3: Normal-Low; 2: Low; 1: Very low.

This test yields a general or overall score for the social and personal competencies studied and the six dimensions or subscales mentioned, which are closely interrelated.

Conventional statistical methods were used to calculate the norms for classifying the behavioral profiles that the scale can measure. The scale was normalized for more than 3,790 participants in the Household Survey conducted by the IDB in the Dominican Republic from late 2010 to early 2011.

A high score on the overall scale is interpreted as a high level of development achieved by the individual in terms of social and personal competencies.

9.8. Basic requirements for administration

The CPS is a battery for groups or individuals. It is easily administered simultaneously to groups of 30 or 40, as long as appropriate rooms are available and the physical and psychological conditions necessary for its administration are present.

There is no time limit, but a person with average reading skills can answer all the questions in 15 to 23 minutes. Populations with very low educational levels may take longer. In the case of people with serious reading deficiencies, the test can be read to the individual like a survey, but this mode of administration has not yet been validated.

9.9. Relevance for Latin America

This instrument was specifically designed for a vulnerable population in Latin America. Thus, the questions have been tailored to a young population with an average reading level, and the test time has been reduced to a maximum of 23 minutes.

9.10. Administration
This test was designed for the evaluation of the Youth and Employment Program and was used in several phases of the evaluation. See Ibarrarán, Rosas (2009) and Ibarrarán et al. (2014).

9.11. Considerations for users and comments

- The CPS is a short, easy, and flexible battery that can be administered to groups or individuals.
- Like the tests developed in *Disconnected*, it can be used without participating in a global study.

10. Relevant links

- **Workforce connections**

A study developed by Child Trends through the Workforce Connections Project, managed by FHI360 and financed by the USAID Office of Education. It reviews more than 384 definitions of socioemotional skills that are key to the success of young adults in the labor market.

http://www.childtrends.org/wp-content/uploads/2015/06/2015-24AWFCSoftSkillsExecSum.pdf

- **Other references on skills measurement for young adults**

    o ToolFind, United Way of Mass Bay with NIOST

        www.toolfind.org

    o Compendium of Assessment and Research Tools (CART), RMC Research Corporation

        http://cart.rmcdenver.com

    o Measurement Tools for Evaluating Out-of-School Time Programs, Harvard Family Research Project

        www.hfrp.org/out-of-school-time/publications-resources

    o Assessment Tools in Informal Science, PEAR at Harvard University, in collaboration with 4-H

    www.pearweb.org/atis

    o Supporting Evaluation and Research Capacity Hub website, CYFAR/USDA

    https://cyfernetsearch.org/

    o Compendium of Student, Teacher, and Classroom Measures Used in NCEE Evaluations of Educational Interventions, IES and Mathematica

        http://ies.ed.gov/ncee/pubs/20104012/pdf/20104013.pdf

    o Online Evaluation Resource Library (OERL), SRI International

        http://oerl.sri.com

- o   Youth Development Outcomes Compendium, Child Trends

  http://childtrends.org/?publications=youth-development-outcomes-compendium

- o   Compendium of Preschool - Elementary School SEL and Associated Assessment Measures, CASEL

  http://www.casel.org/library/2013/11/4/compendium-of-preschool-through-elementary-school-socialemotional-learning-and-associated-assessment-measures

- o   Afterschool Youth Outcomes Inventory, PASE

  http://www.pasesetter.com/outcomes/outcomes_inventory.html

- o   SEL Measures for Middle School Youth, UW Social Development Research Group for Raikes Foundation

  http://www.raikesfoundation.org/Documents/SELTools.pdf

- o   Measuring Student Engagement in Upper Elementary Through High School, REL Southeast

  http://ies.ed.gov/ncee/edlabs/regions/southeast/pdf/REL_2011098.pdf

- **Cognitive Atlas:**

A project led by Russell Poldrack, Professor of Psychology at Stanford University, in collaboration with the University of California, Los Angeles (UCLA) through the UCLA Center for Computational Biology and the UCLA Consortium for Neuropsychiatric Phonemics. The purpose of this project is to develop a knowledge base for sharing and characterizing the state of the art in cognitive science research.

The atlas presents a clear characterization of different skills dimensions and a list of survey batteries for assessing them. The initiative derives from a consideration of existing mental processes and how they relate to the tasks and instruments used to assess them.

http://www.cognitiveatlas.org

http://journal.frontiersin.org/Journal/10.3389/fninf.2011.00017/full

http://www.cognitiveatlas.org/task-collection/id/tco_50ec58d050cf5

- Inventory of Spanish translations of English test names:

http://www.medtrad.org/panacea/IndiceGeneral/n2_PsicoscalesMartinJacod.pdf

Part 3: Skills Usage/Demand Surveys

11.   The UK Commission's Employer Skills Survey UKCES-ESS

11.1.   Organization responsible for development and/or execution:

UK Commission for Employment and Skills.
https://www.gov.uk/government/publications/ukces-employer-skills-survey-2013
https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/525444/UKCESS_2015_Report_for_web__May_.pdf

11.2.   General Objective

Provide information about the skills required by employers and their investments in training to determine their short-term demand for skills and current and future demand for training.

The survey explores the challenges that employers face in terms of the skills of their current employees or those required of new hires. It also provides information about the level and nature of investments in training and the recruitment of young adults and graduates. In short, it looks at the relationship between skills, training, and employer business strategies. The UK uses this survey as part of a labor-market intelligence-gathering system to anticipate skills demand in the productive sector.[30]

11.3.   History
This survey is part of a long tradition of labor market data collection and is the second version of an effort that began in 2011 to homogeneously collect information in the four countries of the United Kingdom (several isolated surveys had been conducted since the 1990s in each of the countries). The survey is conducted every two years and supplemented with a survey that explores how companies interact with the labor market to meet their skill and competency needs.[31] The latest version of this survey was conducted in 2015.

11.4.   Target population

---

[30] For more details on the use of this survey to anticipate skills demand, see Gonzalez-Velosa and Rucci (2016). This report also mentions a CEDEFOOP field survey on skills needs in the European Union that collects more detailed information on generic and job-specific skills.

[31] This survey is called The UK Commission's Employer Perspectives Survey" (UKCEPS). The UKCESS identifies employer demand for skills, competencies, and training, and the UKCEPS analyzes how employers meet this demand.

This survey interviews employers in public and private business establishments. Note that the emphasis is on the business establishment and not the enterprise; this is important, because enterprises can have several business establishments with different needs. Sole traders and establishments with just one employee and no working proprietors were excluded.

The survey was conducted across the entire United Kingdom: England, Northern Ireland, Scotland, and Wales.

11.5.   Type of skills

The survey directly asks employers what skills are hard to find when recruiting new personnel. While there is no direct evaluation, employers are queried about the following skills: basic computer literacy, advanced IT or software skills, communication (oral and written), language, customer relations, teamwork, problem solving, planning and organization, strategic management, numeracy, literacy, and technical, practical, or job-specific skills.

11.6.   Content

The core survey of establishments (22 minutes) and the Investment in Training follow-up survey (for a subset of establishments that reported having provided training in the previous 12 months and agreed to participate).

Core survey: Seven sections: 1) general information about the establishment; 2) recruitment; this section specifically asks whether lack of the skills listed in the preceding section have been an obstacle to hiring; 3) skills demand and perceptions of skills shortages, which includes expressions of employer satisfaction with workers in eight occupations and specific skills shortages in two randomly selected occupations; 4) skills generation and training needs, by occupation; 5) workforce development and training, which includes length of the training in number of days, number of workers who received training by type of occupation, and the relation to national certification; 6) skills usage, consisting basically of questions about "high performance working practices"[32] and 7) the establishment's business strategy and structure.

11.7.   Evaluation

The survey has no tests to evaluate, so no evaluation is conducted. Its results are available to the public, and reports on specific topics are also published.

---

[32] Human resource practices aimed at stimulating worker performance, increasing worker engagement and commitment, and substituting the conventional hierarchical organizational structure with a more participatory one.

11.8. Basic requirements for administration

Few details are provided in this regard, but it is known that some 91,279 telephone interviews were conducted (sampling by establishment and maximum quotas per country). Four reputable agencies were contracted to conduct the survey (BMG, IFFIPSOS MORI, Gaff NOP).

11.9. Relevance for Latin America

This survey is specific to the United Kingdom and is therefore not intended for use in other countries. However, at least three characteristics are relevant for the region.

- The effort to gather information that is detailed, consistent, and comparable over time and among countries, as well as the efforts to use the survey results to evaluate government investment in skills development and guide employer and worker decision-making.
- It is also an important example of the type of questions, length, and methodology to use in surveying employers. In fact, this survey has been used as a reference for other skills usage and demand surveys in the region (for example, the STEP and Disconnected employer surveys).
- It can also serve as an example of estimations of training costs in business establishments. See: https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/303497/ukces-employer-skills-survey-13-technical-report.pdf.

11.10. Administration

The survey's ultimate purpose is to gather information for an updated report and develop policies and practices that will guarantee the effectiveness of skills investments. This information is tabulated and disseminated to the public, because it is intended to inform not only the government but employers and workers as well, so that they can make informed decisions.

A recent list of UKCES publications based on the surveys can be found at:

https://www.gov.uk/government/organisations/uk-commission-for-employment-and-skills/about/research

11.11. Considerations for users and comments

This survey, like the other employer surveys on skills demand, is very useful for current and short-term diagnostic studies. The basic criteria for its use are that it be comparable over time, conducted regularly and predictably, and that the information be properly disseminated.

If the purpose is to forecast future skills demand, it is important to combine these surveys with other, more quantitative methodologies, since there may be inconsistencies between sectors, and employers' responses are highly subjective and situation-dependent[33].

12. Talent Shortage Survey

12.1. Organization responsible for development and/or execution

ManpowerGroup

http://manpowergroup.us/talent-shortage

12.2. General Objective

This is a perception survey designed to determine whether employers are facing talent shortages (difficulty finding employees to fill vacancies), the hardest type of job to find good candidates for, the reason for this problem, the impact of this talent shortage on businesses and their ability to meet customer demands, and employer strategies to solve this problem.

12.3. History
Annual survey launched in 2006, the most recent being the 2015 survey.

12.4. Target population
Unit of analysis: employers.

The last survey interviewed over 41,000 employers in 42 countries (for the 2013 survey, the figure was over 38,000 employers in 42 countries):
Peru, Argentina, Brazil, Panama, Colombia, Costa Rica, Guatemala, Mexico, Japan, India, Turkey, New Zealand, Hong Kong, Israel, Hungary, Taiwan, Bulgaria, Australia, Austria, Germany, Greece, Romania, the United States, Italy, Finland, Poland, Sweden, Switzerland, Canada, China, Slovakia, France, Norway, Slovenia, Belgium, United Kingdom, Czech Republic, Singapore, South Africa, the Netherlands, Spain, Ireland.

12.5. Type of skills
Since this is perception survey, no questions are asked about specific skills; they are presented as a response option for the question about why it is hard to find good candidates to fill vacancies. In this

---

[33] Gonzalez-Velosa and Rucci (2016) describe the different method used to anticipate skills demand.

context, the information about skills is very different from that obtained from surveys of individuals, since there is no direct measurement and it is based on employer perceptions rather than employee performance.

The technical competencies (hard skills) and workplace competencies (soft skills) options differ. The other options offered are: lack of applicants, lack of experience, applicant was looking for higher pay than offered, applicant was unwilling to work in the geographic location offered or move, poor image of the sector or occupation, applicant was unwilling to work half-time, poor image of the company or its culture, and over-qualified applicants.

12.6. Content

- Difficulty finding employees to fill vacancies: open question. 36% on average (16,272 employers)
- Type of job for which it is hardest to find good candidates: how many categories? Report shows the top 10 (Skilled trade workers, engineers, technicians, sales representatives, accounting and financial staff, executives and directors, sales directors, IT staff, support staff, drivers.)
- The reason for this difficulty: Lack of technical competencies (hard skills), lack of workplace skills (soft skills), lack of applicants, lack of experience, applicant expected higher pay than offered, applicant was unwilling to work in the geographic location offered or move, poor image of the sector or occupation, applicant was unwilling to work half-time, poor image of the company or its culture, over-qualified applicants.
- Impact of this talent shortage on the company:
  - Level of skill in the company to meet customer demands: high, medium, low, no impact, don't know (What does "high" mean? Is this an open question that employers subjectively answer?).
  - Mechanisms: reducing the ability to serve customers, reducing the level of competitiveness/productivity, increasing worker turnover, reducing innovation and creativity, lowering worker morale and commitment, increasing the cost of benefits.

- Employer strategies to surmount these difficulties:

  - Personnel practices: provide training to current personnel, adopt new or non-traditional company recruitment practices, create temporary positions for talented current employees who have skills that are in demand (chiefly executives and senior staff), raise salaries, offer applicants clear professional development opportunities, offer more flexible employment conditions.
  - Sources of talent: adapting sources of talent to include groups with untapped talents. The options for this question are: recruit outside the region; recruit candidates from other countries, youth, women, older persons, and veterans; recruit personnel that lack the required skills but have growth potential; forge ties with educational institutions to develop programs aligned with company needs, consider opening new plants/offices or expanding existing ones in areas where the talent is located.
  - Work models: heighten the focus on improving the company's talent flow, change existing work models, include temporary workers in the process, offer applicants virtual work options.

o   No strategy.

12.7.   Evaluation – scoring and interpretation

Tabulation of answers and comparison of percentages.

12.8.   Basic requirements for administration
This survey is not for conducting but consulting.

12.9.   Relevance for Latin America
It was conducted in eight countries in the region: Peru, Argentina, Brazil, Panama, Colombia, Costa Rica, Guatemala, and Mexico.

12.10.   Administration
The survey is constantly cited as a reference on the situation reported by employers about the difficulty finding workers who meet the requirements of the available jobs in the company.

12.11.   Considerations for users and comments
- The self-reporting of "Talent Shortage" is endogenous. Countries that grow faster not only have higher demands but change more rapidly over time.

- It is by no means a definitive indicator of the skills gap or mismatch between the supply and demand for skills. It is simply employer perceptions, supplemented with other measurements, such as the low level of workforce training. It should be borne in mind that this perception of the skills shortage may be the result of information and intermediation problems in some businesses, industries, occupations, or countries.

13.   EWCS-European Working Conditions Survey

13.1.   Organization responsible for development and/or execution

European Foundation for the Improvement of Living and Working Conditions (Eurofound) - European Working Conditions Observatory.

13.2. General Objective

The General Objective is to provide harmonized information on working conditions in Europe. Specific objectives: 1. Assess and quantify the working conditions of employees and the self-employed across Europe, 2. Analyze relationships between different aspects of working conditions, 3. Identify groups at risk and issues of concern as well as progress, 4. Monitor trends by providing homogenous indicators on these issues, 5. Contribute to policy development in the region.

13.3. History

Five previous rounds: 1990-91; 1995-96; 2000; 2005, 2010. The latest version includes 28 EU Member States, Norway, Sweden, the former Yugoslav Republic of Macedonia, Montenegro, Serbia, and Turkey.

13.4. Target population

Employers and workers.
All residents of the participating countries over the age of 15 or 16 that were working at the time of the survey (at least 1 hour the week prior to the interview, according to the ILO definition).

13.5. Type of skills

Skills are not measured directly. Instead, questions are asked to determine tasks and skills usage on the job. The survey is a bit more focused on identifying the nature and conditions of the work than on skills usage per se. It provides the following classification of questions on cognitive, psychosocial, and physical factors. However, other categories, such as working conditions, contain questions associated with all categories.

- Cognitive factors: 9 questions, 4 to determine the use of computers, the Internet, or e-mail for work purposes, and the others related to quality standards, whether workers are required to meet them, and who evaluates the quality of the work; whether workers learn new skills on the job; whether they use their own ideas; whether the work requires complex tasks (definition-open question); task rotation.

- Psychosocial factors: 7 questions on working with customers, dealing with angry customers, knowing what is expected from their work, performing tasks that conflict with their personal values, needing to hide their feelings, the consequences of mistakes in their work.

- Physical factors: 12 questions: exposure to vibration from manual tools or machinery, repetitive hand or arm movements, lifting heavy loads, and other occupational risks.

General characteristics of the work, hours, intensity, health and well-being, social relationships, job satisfaction, organization of the work (task rotation, teamwork, innovation, multitasking). Development module (access to training activities, who pays for them, and whether the skills match work requirements (self-assessment).

The survey is divided into four groups of occupations (based on their ISCO classification): A) Highly skilled administrative personnel: 1. Legislators, senior officials, and managers; 2. Professionals; B) Less skilled administrative personnel; 3. Technicians and associate professionals; 4. Clerical support workers; 5. Personal services workers and sales workers; C) Skilled workers: 6. Skilled agricultural and fishery workers; 7. Craft and related trades workers. D) Less-skilled workers: 8. Plant and Machine Operators and Assemblers; 9. Elementary occupations; 10. Armed forces occupations.

13.6. Content

General: employment status, working hours and schedule, organization of the work, learning and training, physical and psychosocial risk factors, health and safety, work-life balance, worker participation, earnings and financial security.
Quality of the work and job, health and well-being in the workplace, organization of the work, working conditions by sector, employability and job security, working hours and work-family balance, gender issues.

13.7. Evaluation

Many activity questions use a scale of 1 to 7, where 1 is always and 7 never. The rest are yes/no questions.

13.8. Basic requirements for administration

Personal surveys in the home or outside the workplace (44 minutes)

13.9. Relevance for Latin America
N/A

13.10. Administration
N/A

13.11.   Considerations for users and comments

This is an example of an employer survey.

III. References

Autor, David H. and Michael J. Handel "Putting Tasks to the Test: Human Capital, Job Tasks, and Wages." Journal of Labor Economics, Vol. 31, No. 2, The Princeton Data Improvement Initiative (Part 2, April 2013), pp. S59-S96.

Bassi, Marina, Matias Busso, Sergio Urzúa, and Jaime Vargas (2012) "Disconnected: Skills, Education, and Employment in Latin America." Inter-American Development Bank. Social Sector. Division of Education. https://publications.iadb.org/handle/11319/427?locale-attribute=en

Bodewig, Christian, and Reena Badiani-Magnusson. Vietnam Development Report 2014 - Skilling up Vietnam: Preparing the workforce for a modern market economy. Washington, D.C., USA: The World Bank, 2013.

Brea, Mayra (2011) "Informe de la Interpretacion de las escalas CPS, Rosenberg y Grit y porpuestas de revision para mejorar confiabilidad" unpublished. http://documentos.bancomundial.org/curated/es/418161467995809251/pdf/105233-SPANISH-WP-P096605-PUBLIC-A2I-request-ACS.pdf

Card, David, Pablo Ibarrarán, Ferdinando Regalia, David Rosas, and Yuri Soares, 2007. "The Labor Market Impacts of Youth Training in the Dominican Republic: Evidence from a Randomized Evaluation." NBER Working Papers 12883, National Bureau of Economic Research, Inc.

Carroll, J B 1993. "Human Cognitive Abilities: A Survey of Factor-Analytic Studies." New York: Cambridge University Press.

Cloninger, Przybeck, Svrakic, and Wetzel (1994) "The Temperament and Character Inventory (TCI): A Guide to its Development and Use" Center for Well-Being. Department of Psychiatry, Campus Box 8134, Washington University School of Medicine, 660 S. Euclid Ave., St. Louis, MO 63110.

Cronbach, Lee J. (1951) "Coefficient alpha and the internal structure of tests" Psychometrika. 16 (3): 297–334.

Costa, P. T and R.R. McCrae (1992) "Four Ways Five Factors are Basic." Personality and Individual Difference 13(6), 653-665.

Del Carpio, Ximena, Yuki Ikeda, and Michele Zini. "Lao PDR - Skills for quality jobs and development in Lao PDR: a technical assessment of the current context." (2013).

Desjardins, R. and K. Rubenson (2011), "An Analysis of Skill Mismatch Using Direct Measures of Skills", OECD Education Working Papers, No. 63, OECD Publishing. doi: 10.1787/5kg3nh9h52g5-en

Desjardins, R. and A. Warnke (2012), "Ageing and Skills: A Review and Analysis of Skill Gain and Skill Loss Over the Lifespan and Over Time", OECD Education Working Papers, No. 72, OECD Publishing. doi: 10.1787/5k9csvw87ckh-en

Dundar, Halil, Benoit Millot, Yevgeniya Savchenko, Harsha Anturupane, and Tillakaratne A. Piyasiri. Building the skills for economic growth and competitiveness in Sri Lanka. Washington, D.C., USA: The World Bank, 2014.

Duckworth, A. L. and Quinn, P. D. (2009). Development and validation of the Short Grit Scale (Grit-S). Journal of Personality Assessment, 91, 166-174. (pdf)

Duryea, S. and A. Williamson (2016) "Indicators for Measuring Psycho-Social Competences in Youth: A Landscape Analysis." Mimeo.

Felstead, A., Gallie, D., Green, F., and Zhou, Y. (2007). Skills at Work, 1986 to 2006. Oxford: ESRC Centre on Skills, Knowledge and Organisational Performance.

George, D. and Mallery, P. (2003). SPSS for Windows step by step: A Simple Guide and Reference. 11.0 Update (4th ed.). Boston: Allyn & Bacon.

Goldberg, L. R. (1999). A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models. In I. Mervielde, I. Deary, F. De Fruyt, and F. Ostendorf (Eds.), Personality Psychology in Europe, Vol. 7 (pp. 7-28). Tilburg, The Netherlands: Tilburg University Press.

Goldberg, L. R., Johnson, J. A., Eber, H. W., Hogan, R., Ashton, M. C., Cloninger, C. R., and Gough, H. C. (2006). The International Personality Item Pool and the future of public-domain personality measures. Journal of Research in Personality, 40, 84-96.

Gonzalez-Velosa, Carolina and G. Rucci (2016) Methods to Anticipate Skills Demand. IDB Technical Note IDB-TN-954. February.

Gosling, S. D., Rentfrow, P. J., and Swann, W. B., Jr. (2003). A Very Brief Measure of the Big Five Personality Domains. *Journal of Research in Personality, 37, 504-528.*

Guerra, N., Modecki, K., and Cunningham, W. (2014). Developing social-emotional Skills for the labor market: The PRACTICE model. World Bank Policy Research Working Paper (7123).

Guvenen, Fatih; Burhan Kuruscu; Satoshi Tanaka, and David Wiczer (2015) "Multidimensional Skill Mismatch." NBER Working Paper No. 21376

Handel, Michael (2008) "Measuring Job Content: Skills, Technology and Management Practices" Institute for Research on Poverty, University of Wisconsin, Madison, WI. Discussion Paper No. 1357-08. Available at: http://www.irp.wisc.edu/publications/dps/pdfs/dp135708.pdf

Handel, Michael (2010) "What Do People Do at Work? A Profile of U.S. Jobs from the Survey of Workplace Skills, Technology, and Management Practices (STAMP)." (OECD, forthcoming).

Hanushek, Schwerdt, Wiederhold, and Woessmann (2013) "Returns to Skills around the World: Evidence from PIAAC," OECD Education Working Papers, No. 101, OECD Publishing.

Heckman, James., Jora Stixrud, and Sergio Urzúa (2006) "The Effects of Cognitive and Noncognitive Abilities on Labor Market Outcomes and Social Behavior." Journal of Labor Economics 24 (3), 411-482.

Heckman, James, and Tim Kautz (2013) "Fostering and Measuring Skills: Interventions That Improve Character and Cognition." NBER Working Paper No. 19656.

Hogarth Terence (2016) "Designing an Employers Skills Survey: notes on how to develop a survey to meet a range of policy issues relating to the demand for, and the supply of skills." Mimeo.

Ibarrarán Pablo and David Rosas Shady (2009) "Evaluating the impact of job training programmes in Latin America: evidence from IDB funded operations." Journal of Development Effectiveness, Taylor & Francis Journals, vol. 1(2), pp. 195-216.

Ibarrarán, Pablo, Ripani, Laura, Taboada, Bibiana Villa, Juan Miguel, and García, Brígida (2012) "Life Skills, Employability and Training for Disadvantaged Youth: Evidence from a Randomized Evaluation Design." IZA Discussion Papers 6617, Institute for the Study of Labor (IZA).

Ibarrarán, Pablo, Jochen Kluve, Laura Ripani, and David Rosas Shady (2015) "Experimental Evidence on the Long-Term Impacts of a Youth Training Program." IZA Discussion Papers 9136, Institute for the Study of Labor (IZA).

John, O. P. and Srivastava, S. (1999). The Big-Five trait taxonomy: History, measurement, and theoretical perspectives. In L. A. Pervin and O. P. John (Eds.), Handbook of personality: Theory and research (Vol. 2, pp. 102–138). New York: Guilford Press.

Kim, HoonHo and Lalancette, Diane (2013) "Literature Review on the Value-Added Measurement in Higher Education" OECD Publishing. Available at: http://www.oecd.org/edu/skills-beyond-school/Litterature%20Review%20VAM.pdf

Liang, Xiaoyan, and Shuang Chen. Developing skills for economic transformation and social harmony in China: study of Yunnan Province. Washington, D.C., USA: The World Bank, 2013.

Lippman, Laura, Renee Ryberg, Rachel Carney, and Kristin Anderson Moore (2015) "Key "Soft Skills" that Foster Youth Workforce Success: Toward a Consensus across Fields." http://www.childtrends.org/?publications=key-soft-skills-that-foster-youth-workforce-success-toward-a-consensus-across-fields-executive-summary

Markey, P. M. and Markey, C. N. (2009). A brief assessment of the interpersonal circumplex: The IPIP-IPC. Assessment. 16, 352-361

OECD (2013), OECD Skills Outlook 2013: First Results from the Survey of Adult Skills. OECD Publishing. Available at: http://dx.doi.org/10.1787/9789264204256-en

OECD. (2012). Better Skills, Better Jobs, Better Lives: A Strategic Approach to Skills Policies. OECD Publishing.

Pellizzari, M. and A. Fichen (2013), "A New Measure of Skill mismatch: Theory and Evidence from the Survey of Adult Skills (PIAAC)," OECD Social, Employment and Migration Working Papers, No. 153, OECD Publishing. doi: 10.1787/5k3tpt04lcnt-en.

Pierre, G., Maria Laura Sanchez Puerta, Alexandria Valerio, and Tania Rajadel (2014) "STEP Skills Measurement Surveys - Innovative Tools for Assessing Skills." Available at: http://microdata.worldbank.org/index.php/catalog/2010/download/30655.

Renau, V., Oberst, U., Gosling, S. D., Rusiñol, J., and Chamarro, A (2013). Translation and validation of the Ten-Item-Personality Inventory into Spanish and Catalan. *Aloma.Revista de Psicologia, Ciències de l'Educació i de l'Esport, 31,* 85-97.

Rosenberg, M.: Society and the adolescent self-image. Society of the adolescent self-image 1965; 11(2)

Rosenberg, M.: Society and the adolescent self-image (Revised edition). Middletown, C. T.: Wesleyan University Press 1989

Rotter, Julian B. 1966. Generalized Expectancies for Internal versus External Control of Reinforcement. Washington, D.C.: American Psychological Association.

Martín Albo, J., Núñez, J. L., Navarro, J. G., and Grijalvo, F.: The Rosenberg Self-Esteem Scale: Translation and Validation in University Students. The Spanish Journal of Psychology 2007; 10(2):458–467

Saucier, G. (1997). Effects of variable selection on the factor structure of person descriptors. Journal of Personality and Social Psychology, 73, 1296-1312.

Streiner, D.L. Starting at the beginning: an introduction to coefficient alpha and internal consistency. J Pers Assess. 2003;80:99-103

Valerio, Alexandria, Maria Laura Sanchez Puerta, Gaëlle Pierre, Tania Rajadel, and Sebastian Monroy Taborda. STEP Skills Measurement Program - Snapshot 2014. Washington, D.C., USA: The World Bank, 2014.

WEF (2014) "Matching Skills and Labour Market Needs Building Social Partnerships for Better Skills and Better Jobs." Global Agenda Council on Employment World Economic Forum.

Wilson-Ahlstrom et al. (2014) "From Soft Skills to Hard Data: Measuring Youth Program Outcomes" Published by The Forum for Youth Investment, January. Available at: http://forumfyi.org/files/soft_skills_hard_data_0.pdf

World Bank (2012) "World Development Report 2013: Jobs" World Bank Publications, 2012.

IV.   Appendix – Test Performance Scales

A.   PIAAC- Literacy: Description of reading proficiency levels with their respective cut scores

Level	Types of tasks successfully completed at each level of performance
Below Level 1 *Less than 176*	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 asked to locate information in brief continuous texts. However, in this case, the information can be located as if the text were non-continuous. 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 at this level do not make use of any features specific to digital texts.
1 *176-225*	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 piece of information that is identical or synonymous with the information given in the question or statement. Some tasks, as in the case of non-continuous texts, may require the respondent to enter personal information into a document. Competing information may be present, but very little. Some tasks may require matching different pieces of information. The respondent is expected to have knowledge and skill in recognizing basic vocabulary in order to understand the meaning of the texts.
2 *226-275*	At this level, the medium of the texts may be digital or printed, and the texts may be comprised of continuous, non-continuous, or mixed types. Tasks at this level require the respondent 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 • integrate two or more pieces of information based on criteria • compare and contrast or reason about the information requested in the question • navigate within digital texts to access and identify information from various parts of a document
3 *276-325*	At this level, texts are often dense or lengthy and include multiple pages of continuous, non-continuous, or mixed text. Understanding text and rhetorical structures, especially when navigating complex digital texts, become central to successfully completing tasks. Tasks require the respondent to identify, interpret, or evaluate one or more pieces of information, and often require varying levels of inferencing. Many tasks require the respondent to construct his response by searching different chunks of text or to perform multistep operations to identify and formulate responses. Often, tasks also demand that the respondent disregard irrelevant or inappropriate information in order to answer accurately. Competing information is often present, but it is not more prominent than the correct information.
4 *326-375*	Tasks at this level often require respondents to perform multistep operations to integrate, interpret, or synthesize information from multiple types of texts–complex or lengthy, continuous, non-continuous or mixed. Complex inferences and applications of background knowledge may be needed to successfully complete the task. Many tasks require identifying and understanding one or more specific, secondary ideas in the text in order to interpret or evaluate subtle evidence or persuasive 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.
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 evidenced-based arguments. Application and evaluation of logical and conceptual models of ideas may be required to accomplish tasks. Evaluating confidence of evidentiary sources and selecting key information is normally a key requirement. Tasks often require the respondent to be aware of subtle rhetorical cues and to make high-level inferences or use specialized background knowledge.

B.   PIAAC- Numeracy: Description of numeracy proficiency levels with their respective cut scores

Level	Types of tasks successfully completed at each level of performance
Below Level 1 *Less than 176*	Tasks at this level require the respondent to perform only simple processes such as counting, sorting, performing basic arithmetic operations with whole numbers of money, or recognizing common spatial representations in common, concrete contexts in which the mathematical content is explicit with no distractors and little or no text.
1 *176-225*	The majority of tasks at this level require the respondent to perform basic mathematical processes in common, concrete contexts where the mathematical content is explicit with minimal text or distractors. Tasks usually require simple processes such as counting, sorting, performing basic arithmetic operations, understanding simple percentages, such as 50%, and locating and identifying elements of simple graphical or spatial representations.
2 *226-275*	At this level, the respondent is required to identify and deal with mathematical information and ideas in a range of common contexts where the mathematical content is visual or explicit with relatively few distractors. Tasks tend to require the application of two or more steps or processes involving decimals of one or two digits, percentages, and fractions; simple measurements and spatial representation; estimation; and the interpretation of relatively simple data and statistics in texts, tables, and graphs.
3 *276-325*	At this level, the respondent is required to understand mathematical information that may not be explicit, embedded in contexts that are not always familiar and represented in more complex ways. These tasks require several steps and may involve problem-solving strategies and relevant processes. Tasks will include the application of numerical concepts and spatial sense; recognizing and working with mathematical relationships, patterns, and proportions expressed both numerically and verbally; and interpretation and basic analysis of data and statistics in texts, tables, and graphs.
4 *326-375*	At this level the respondent is required 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 processes and strategies. Tasks usually require analysis and more complex reasoning about quantities and data; statistics and probability; spatial relationships; proportions; and formulas. At this level, the respondent may be required to understand arguments or formulate well-reasoned explanations for answers or choices.
5 *376-500*	Tasks at this level require the respondent to understand complex representations and abstract and formal mathematical and statistical ideas, possible embedded in complex texts. The respondent may have to integrate multiple types of mathematical information where translation and interpretation is required; draw inferences; develop or work with mathematical models or arguments; and justify, evaluate, and critically reflect on responses.

C.   ALL: Prose, Document Literacy, and Numeracy Proficiency Scales

	Prose	Document	Numeracy
Level 1 (0 to 225)	Most of the tasks in this level require the respondent to read relatively short text to locate a single piece of information which is identical to or synonymous with the information given in the question or directive. If plausible but incorrect information is present in the text, it tends not to be located near the correct information.	Tasks in this level tend to require the respondent either to locate a piece of information based on a literal match or to enter information from personal knowledge onto a document. Little, if any, distracting information is present.	Tasks in this level require the respondent to show an understanding of basic numerical ideas by completing simple tasks in concrete, familiar contexts where the mathematical content is explicit with little text. Tasks consist of simple, one-step operations such as counting, sorting dates, performing simple arithmetic operations or understanding common and simple percents such as 50%.
Level 2 (226 to 275)	Some tasks in this level require respondents to locate a single piece of information in the text; however, several distracters or	Tasks in this level are more varied than those in Level 1. Some require the respondents to match a single piece of	Tasks in this level are fairly simple and relate to identifying and understanding basic mathematical concepts embedded in a range

	plausible but incorrect pieces of information may be present, or low-level inferences may be required. Other tasks require the respondent to integrate two or more pieces of information or to compare and contrast easily identifiable information based on a criterion provided in the question or directive.	information; however, several distracters may be present, or the match may require low-level inferences. Tasks in this level may also ask the respondent to cycle through information in a document or to integrate information from various parts of a document.	of familiar contexts where the mathematical content is quite explicit and visual with few distracters. Tasks tend to include one-step or two-step processes and estimations involving whole numbers, benchmark percents and fractions, interpreting simple graphical or spatial representations, and performing simple measurements.
Level 3 (276 to 325)	Tasks in this level tend to require respondents to make literal or synonymous matches between the text and information given in the task, or to make matches that require low-level inferences. Other tasks ask respondents to integrate information from dense or lengthy text that contains no organizational aids such as headings. Respondents may also be asked to generate a response based on information that can be easily identified in the text. Distracting information is present, but is not located near the correct information.	Some tasks in this level require the respondent to integrate multiple pieces of information from one or more documents. Others ask respondents to cycle through rather complex tables or graphs which contain information that is irrelevant or inappropriate to the task.	Tasks in this level require the respondent to demonstrate understanding of mathematical information represented in a range of different forms, such as in numbers, symbols, maps, graphs, texts, and drawings. Skills required involve number and spatial sense, knowledge of mathematical patterns and relationships and the ability to interpret proportions, data and statistics embedded in relatively simple texts where there may be distracters. Tasks commonly involve undertaking a number of processes to solve problems.
Level 4 (326 to 375)	These tasks require respondents to perform multiple-feature matches and to integrate or synthesize information from complex or lengthy passages. More complex inferences are needed to perform successfully. Conditional information is frequently present in tasks at this level and must be taken into consideration by the respondent.	Tasks in this level, like those at the previous levels, ask respondents to perform multiple-feature matches, cycle through documents, and integrate information; however, they require a greater degree of inferencing. Many of these tasks require respondents to provide numerous responses but do not designate how many responses are needed. Conditional information is also present in the document tasks at this level and must be taken into account by the respondent.	Tasks at this level require respondents to understand a broad range of mathematical information of a more abstract nature represented in diverse ways, including in texts of increasing complexity or in unfamiliar contexts. These tasks involve undertaking multiple steps to find solutions to problems and require more complex reasoning and interpretation skills, including comprehending and working with proportions and formulas or offering explanations for answers.
Level 5 (376 to 500)	Some tasks in this level require the respondent to search for information in dense text which contains a number of plausible distracters. Others ask respondents to make high-level inferences or use specialized background knowledge. Some tasks ask respondents to contrast complex information.	Tasks in this level require the respondent to search through complex displays that contain multiple distracters, to make high-level text-based inferences, and to use specialized knowledge.	Tasks in this level require respondents 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, draw inferences, or generate mathematical justification for answers.

D.   STEP: Proficiency Scales

Direct measurement of reading proficiency | Key indicators

Core Literacy Assessment	
Did not pass (i.e., fewer than 3 correct responses out of 8 items)	Respondent's reading proficiency is below a minimum reading literacy threshold
Pass (i.e., 3 or more correct responses out of 8 items)	Respondent has met a minimum reading literacy threshold

Reading Proficiency Levels

Literacy Below Level 1   **0 to 175**

The tasks at this level require the respondent to read brief texts on familiar topics to locate a single piece of specific information. 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. There is seldom any competing information in the text and the requested information is identical in form to information in the question or directive. While the texts can be continuous, the information can be located as if the text were noncontinuous. Tasks below Level 1 do not make use of any features specific to digital texts.

Literacy Level 1   **176 to 225**

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 which is identical to or synonymous with the information given in the question or directive. Some tasks may require the respondent to enter personal information into a document, in the case of some noncontinuous texts. 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, evaluating the meaning of sentences, and reading of paragraph text is expected.

Literacy Level 2   **226 to 275**

At this level, the complexity of text increases. The medium of texts may be digital or printed, and texts may comprise continuous, noncontinuous or mixed types. Tasks in this level require respondents to make matches between the text and information, and may require paraphrase 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.

Literacy Level 3   **276 to 325**

Texts at this level are often dense or lengthy, including continuous, noncontinuous, mixed or multiple pages. Understanding text and rhetorical structures become more central to successfully completing tasks, especially in navigation of complex digital texts. Tasks require the respondent to identify, interpret or evaluate one or more pieces of information and often require varying levels of inferencing. Many tasks require the respondent construct meaning across larger chunks of text or perform multistep operations in order to identify and formulate responses. Often tasks also demand that the respondent disregard irrelevant or inappropriate text content to answer accurately. Competing information is often present, but it is not more prominent than the correct information.

Literacy Level 4   **326 to 375**

Tasks at this level often require respondents to perform multiple-step operations to integrate, interpret, or synthesize information from complex or lengthy continuous, noncontinuous, mixed, or multiple type texts. Complex inferences and application of background knowledge may be needed to perform successfully. Many tasks require identifying and understanding one or more specific, noncentral ideas 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.

Literacy Level 5   **376 to 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.

Indirect measurement of cognitive skills | Key indicators

Use of reading and writing skills		Intensity of use	Level
Does not do read/write	=	Does not use	0
Read/write documents of 5 pages or less	=	Low	1
Read/write documents of 6 to 25 pages	=	Medium	2
Read/write documents of more than 25 pages	=	High	3

Use of numeracy skills		Complexity of use	Level
Does no math	=	Does not use	0
Measures or estimates sizes, weights, distances, calculate prices or costs, perform any other multiplication or division	=	Low	1
Uses or calculates fractions, decimals or percentages	=	Medium	2
Uses more advanced math such as algebra, geometry, trigonometry	=	High	3

Socio-emotional skills | Items

Socio-emotional Skill	Items
Openness	- Do you come up with ideas other people haven't thought of before? - Are you very interested in learning new things? - Do you enjoy beautiful things, like nature, art and music?
Conscientiousness	- When doing a task, are you very careful? - Do you prefer relaxation more than hard work? - Do you work very well and quickly?
Extraversion	- Do you come up with ideas other people haven't thought of before? - Are you very interested in learning new things? - Are you outgoing and sociable, for example, do you make friends very easily?

Agreeableness	- Do you forgive other people easily? - Are you very polite to other people? - Are you generous to other people with your time or money?
Emotional Stability (Neuroticism)	- Are you relaxed during stressful situations? - Do you tend to worry? - Do you get nervous easily?
Grit	- Do you finish whatever you begin? - Do you work very hard? For example, do you keep working when others stop to take a break? - Do you enjoy working on things that take a very long time (at least several months) to complete?
Hostile Bias	- Do people take advantage of you? - Are people mean/not nice to you?
Decision-making	- Do you think about how the things you do will affect you in the future? - Do you think carefully before you make an important decision? - Do you ask for help when you don't understand something?

Job-relevant skills | Items

Computer use		Intensity of use	Level
"As a part of your work do you use a computer?" "As a part of your life [outside work] have you used a computer in the past 3 months?"			
Does not use a computer/use a computer almost never	=	Does not use	0
Uses computer less than three times per week	=	Low	1
Uses computer three times or more per week	=	Medium	2
Uses computer every day	=	High	3

Contact with clients		Intensity of use	Level
"As part of this work, do you have any contact with people other than co-workers, for example customers, clients, students, or the public?" *			
Does not have any contacts with clients	=	Does not use	0
Involvement scale ranges from 1 to 4	=	Low	1
Involvement scale ranges from 5 to 7	=	Medium	2
Involvement scale ranges from 8 to 10	=	High	3

Solving and learning at work	Intensity of use	Level

Item 1. "Some tasks are pretty easy and can be done right away or after getting a little help from others. Other tasks require more thinking to figure out how they should be done. As part of this work, how often do you have to undertake tasks that require at least 30 minutes of thinking?"			
Never	=	Does not use	0
Less than once per month	=	Low	1
Less than once a week but at least once a month OR at least once a week but not every day	=	Medium	2
Every day	=	High	3
Item 2. "How often does (did) this work involve learning new things?"			
Rarely	=	Does not use	0
At least 2-3 months or at least once a month	=	Low	1
At least once a week	=	Medium	2
Every day	=	High	3

Average of 2 items

Autonomy and repetitiveness	Intensity of use	Level

Item 1. "Still thinking of your work, how much freedom do you have to decide how to do your work in your own way, rather than following a fixed procedure or a supervisor's instructions? Use any number from 1 to 10 where 1 is no freedom and 10 is complete freedom."			
Decision freedom scale from 1 to 2	=	Close to none	0
Decision freedom scale from 3 to 6	=	Low	1
Decision freedom scale from 7 to 9	=	Medium	2
Decision freedom scale 10	=	High	3
Item 2. "How often does (did) this work involve carrying out short, repetitive tasks?"			
Almost all the time	=	Close to none	3
More than half the time	=	Low	2
Less than half the time	=	Medium	1
Almost never	=	High	0

Average of 2 items

E. CPS: TABLE OF STEM NORMS AND CUTS

Meaning	Sten	CPS - Total	CPS-01 Leadership	CPS-02 Behavior in situations of conflict	CPS-03 Self-esteem	CPS-04 Ability to relate to others	CPS-05 Organization-Order	CPS-06 Empathy and Communication Skills
Very high	10	132	21	30	12	15	18	36
High	9	124-131	21	29	12	15	18	34-35
Normal-High	8	116-123	20	27-28	12	14	17	31-33
Normal	7	108-115	18-19	25-26	11	13	16	28-30
	6	99-107	16-17	22-24	10	12	14-15	26-27
	5	91-98	14-15	20-21	9	10-11	13	23-25
	4	83-90	12-13	18-19	8	9	11-12	20-22
Normal-Low	3	75-82	10-11	15-17	7	8	9-10	18-19
Low	2	66-74	9	13-14	6	6-7	8	15-17
Very low	1	0-65	0-8	0-12	0-5	0-5	0-7	0-14

CPS Statistics

	CPS – Total	CPS-01 Leadership	CPS-02 Behavior in situations of conflict	CPS-03 Self-esteem	CPS-04 Ability to relate to others	CPS-05 Organization-Order	CPS-06 Empathy and Communication skills
Average	99.417	15.936	22.302	9.719	11.679	14.179	25.601
Deviation	16.475	3.662	4.683	2.07	2.758	3.138	5.362
Maximum	132	21	30	12	15	18	36
Minimum	0	0	0	0	0	0	0
Cases	3803	3803	3803	3803	3803	3803	3803

V. Appendix: Test and Measurement Design Procedure

1. Operative definition of the term/skill: the limits of those measured and especially those NOT measured are established.

2. Organization of the domain: how the activities and measurements will be presented to the final consumers.

An important part of this step is to use factor analysis to determine how many dimensions are captured in a battery of questions.

Example: In the NAEP Young Adult Literacy Survey (Kirsh and Jungeblut, 1986), the authors conducted a factor analysis and found at least three necessary factors for explaining the variance in the responses to the questions. Thus, the activities with a higher coefficient in the first factor (orthogonal) were those related to prose literacy; in the second factor, document literacy; and in the third factor, the activities required the application of arithmetic operations. This is how the items were defined in IALS.

3. Identification of the characteristics of the tasks developed to assess each skill area.
4. Identification and operationalization of variables for constructing the test.
5. Validation of the variables.
6. Creation of a system for interpretation of the results and linkage with the final report.

VI. Appendix: Tables in AnexosInventario. xls

1. "Table A1.InventariosIPIP": List of personality scales and inventories for which IPIP has developed an alternative and made it available. These inventories are commonly used in the literature.
   The Table includes the number of the test, the abbreviation commonly used in the specialized literature to refer to it, a brief description, the date of publication and of the latest versions/revisions, the company that publishes or owns the rights to it, the age ranges,

information about other available versions, the name in Spanish (if there is a consensus on what to call it) and links.

2. "Table A2. Instrumentos": Non-exhaustive list of tests, activities, and instruments for measuring specific skills not included in the inventory.

3. "Table A3.TestEmpresas": Non-exhaustive list of commercial psychometric tests.