A Comprehensive Review and Development of Measurement Options for Essential Skills Initiatives

Phase 1 – Inventory of Measures

September 2018

SRDC 25 SRSA

SOCIAL RESEARCH AND DEMONSTRATION CORPORATION
SOCIÉTÉ DE RECHERCHE SOCIALE APPLIQUÉE

The Social Research and Demonstration Corporation (SRDC) is a non-profit research organization, created specifically to develop, field test, and rigorously evaluate new programs. SRDC's two-part mission is to help policy-makers and practitioners identify policies and programs that improve the well-being of all Canadians, with a special concern for the effects on the disadvantaged, and to raise the standards of evidence that are used in assessing these policies.

Since its establishment in December 1991, SRDC has conducted over 350 projects and studies for various federal and provincial departments, municipalities, as well as other public and non-profit organizations. SRDC has offices located in Ottawa and Vancouver, and satellite offices in Calgary and Montreal.

Table of contents

Introduction

Overall project objectives

The Office of Literacy and Essential Skills (OLES) currently supports a wide range of Literacy and Essential Skills (LES) projects that include pilots of various training delivery models in both workplace settings and for jobseekers, through the Adult Learning, Literacy, and Essential Skills Program (ALLESP). This program aims to develop, test and evaluate innovative training models that support Canadians in improving their LES, with the aim to help them obtain and keep a job, as well as to adapt and succeed at work. Performance measurement is critical to the achievement of these broader objectives of ALLESP and to the success of the projects it supports.

One of the many challenges is that the measurement options for evaluating success of LES initiatives are vast and complex in nature. The preferred measurement instruments will vary by, among other factors, program objectives, scope of the models, delivery context, and the target population. Measurement options are also quite dynamic with new indicators and evidence on their validity and reliability emerging on a frequent basis not to mention evolution in the delivery models that require new approaches to measurement. The wider LES community would benefit greatly from i) a consolidation of current knowledge on existing measurement options for LES initiatives, as well as ii) the development of a framework for the application of measurement options in different program contexts, with different targeted populations.

SRDC has been contracted by OLES to conduct a comprehensive review of measurement options for assessing outcomes of LES initiatives and to develop a framework to support a broader performance measurement strategy. The project will be conducted in four phases, each with specific objectives:

Phase 1 – Literature review and environmental scan

Through a comprehensive literature review and environmental scan, SRDC will document currently available instruments and measures relevant to all nine Essential Skills, including the core literacy and numeracy skills (Reading, Writing, Document Use, and Numeracy), digital literacy, and the four 'soft' skills that are receiving increasing attention in recent research: Oral Communication, Thinking, Working with Others, and Continuous Learning. As part of this review, SRDC will outline and assess any existing evidence on the validity and reliability of all measures and their suitability for different contexts and populations. The review will also update and expand on the analysis of publicly-available documents by incorporating information form key informant interviews with assessment developers, practitioners, and other key LES stakeholders.

Phase 2 – Framework development

Following the review and scan, SRDC will develop a broader framework to support performance measurement applicable to a range of LES initiatives. This will include a milestones and pathways based approach that incorporates measures linked to both intermediate performance gains and longer-term and employment outcomes, including key contextual variables that may act to create

conditions for success. SRDC will develop the framework by building on the Phase 1 review as well as evidence from earlier SRDC projects such as Pay for Success, Foundations, and UPSKILL.

Phase 3 – Field testing and data analysis

SRDC will undertake additional data analysis and field testing of a selected subset of key skills measures and performance indicators from the Phase 2 framework. This will help determine the statistical properties of these indicators and provide evidence of their suitability as precursors to longer-term outcomes such as employment. This effort will include extended analyses of existing SRDC data sets (e.g., Foundations, Skilling UP, ESSF, and UPSKILL), with the aim of replicating the Pay for Success approach of determining which indicators function best as possible milestones towards longer-term success of LES initiatives. It may also include further field testing of available instruments and measures with active LES initiatives and partners. The scope and focus of field testing will be determined in consultation with ESDC following completion of Phases 1 and 2.

Phase 4 – Final report and recommendations

Once the analysis and field testing phase is complete, SRDC will revise the framework and recommend a series of preferred indicators along with guidelines for their use in evaluating different LES programs in a range of contexts and for various populations. The final report will aim to serve as a practical guide for LES practitioners and policy-makers in selecting suitable indicators for evaluating success of their initiatives. At the same time, it will support OLES in creating alignment and synergies across projects, as part of a broader performance measurement strategy and monitoring of the achievements of ALLESP.

Objectives and structure of this report

This report incorporates the findings of the Phase 1 portion of the project, i.e., a review of available instruments and measures relevant to all nine Essential Skills, as well as any existing evidence on their statistical properties and suitability for different contexts and populations. To supplement existing publically available documentation , we have incorporated into our findings information from key informant interviews with assessment developers and users. In addition, we have undertaken some work to explore measurement options for other soft skills and psychosocial attributes that may be linked with workplace success, but that were developed outside the ES framework.

Background

The framework supporting literacy and essential skills has been utilized for over 20 years, and has resulted in the establishment of ES Profiles for most Canadian occupations, which outline the foundational skills that are needed for success in each occupation. The framework also provides the foundation for most of Canada's National Occupational Standards (NOS) where ES Profiles have been integrated with technical skills requirements in the descriptions of Canadian jobs. In addition, the ES framework is connected to skills measures that have been used in multiple rounds of

international literacy assessment (e.g., IALS in 1994, IALSS in 2003, and PIAAC in 2012), in which Canadian respondents participated.

Literacy and Essential Skills are conceptualized as competencies and capabilities that jobseekers and workers need in order to be successful in the labour markets. LES includes elements of basic literacy, i.e., Reading, Writing, Document Use, Numeracy, and Digital Literacy; as well as non-technical skills such as Thinking, Communication, Working with Others, and Continuous Learning. To be more specific, the conceptualization of the nine Essential Skills are as follows:

1. **Reading Text**: reading materials in the form of sentences or paragraphs

2. **Writing**: Writing text and writing in documents, such as filling in forms, and non-paper-based writing such as typing on a computer

3. **Document Use**: Tasks that involve a variety of information displays in which words, numbers, symbols and other visual characteristics (e.g., lines, colours or shapes) are given meaning by their spatial arrangement

4. **Numeracy**: Using numbers and thinking in quantitative terms to complete tasks

5. **Digital Literacy**: Using different kinds of computer applications and other related technical tools e.g., digital literacy

6. **Thinking Skills**: The process of evaluating ideas or information to reach a rational decisions, e.g., problem solving, decision making, critical thinking, job task planning and organizing, and significant use of memory and finding information

7. **Oral Communication**: Using speech to give and exchange thoughts and information

8. **Working with Others**: Collaboration among individuals and groups e.g., employees working with others to carry out their job tasks in an efficient and effective manner

9. **Continuous Learning**: Participating in an ongoing process of acquiring skills and knowledge e.g., receptivity and engagement in life-long learning.

Researchers and practitioners may use different types of ES assessments depending on participant characteristics and training context. For example, generic assessments that draw upon materials and skill domains used in a variety of work and non-work settings may be appropriate for programs that aim to sharpen transferable Essential Skills to enhance readiness for a wide range of learning and employment contexts. In contrast, assessments that are embedded within or informed by job competency frameworks – i.e., tasks that workers need to carry out efficiently within specific industries – are more suitable for workplace training programs that are more tightly aligned with industry skill needs.

Furthermore, different types of assessments may be integrated into the same project, depending on the comprehensiveness of the program model and the progress its participants make along the employment pathway. For example, a training model for job seekers with multiple barriers can be designed to first strengthen readiness for further learning or employment in general, and then enhance Essential Skills that underlie industry-contextualized job performance. In this case, generic

assessments can be used at the early stage to capture Essential Skills acquisition in general terms, providing an indication of transferable skills participants can potentially bring to further training or to the labour market. Participants who continue with training customized to the skills requirements of a specific sector can have their progress tracked with assessments contextualized to the job performance requirements of that sector.

SRDC's review documents a full range of both of these kinds of assessment tools – generic and industry-contextualized. Within each of these two categories, we further distinguish assessments that use objective competency-based measures (e.g., questions with correct answers) from those that rely on more subjective measures (e.g., self-reports). It is important to note (as we do throughout the report) that even objective assessments can vary in their level of precision, and that measurement error can make some of these tools more suitable for low stakes (e.g., determining progress at a group level) rather than high stakes assessments (e.g., determining individual selection into an education program or job). Finally, based on our experience with other psychosocial measures, we also suggest ways to augment and expand the Essential Skills framework to capture a fuller range of skills and attributes that provide additional information on likelihood of success in today's labour market. We summarize how each of these kinds of assessments have been developed and used in different program contexts and with different target populations, and conclude with a brief summary of how these findings can be used to set the stage for Phase 2 – guidelines for building milestone-based performance measurement frameworks for a wide range of LES initiatives.

Assessments of generic Essential Skills

Objective assessments

Standardized generic assessments have been developed for reading (or prose), document use, numeracy, and to some extent digital literacy. There is no standardized writing assessment for adults, probably because of the difficulties involved in developing objective assessment criteria.

Although the concept of literacy is not new, the first systematic assessment of adult literacy across countries was conducted under the International Adult Literacy Survey (IALS) in 1994. Recent developments in Literacy and Essential Skills research in Canada can be traced back to IALS as the root. Because IALS assessed only individual's proficiency in reading (prose), document use, and numeracy, subsequent Essential Skills research and assessment products tend to focus on these three domains.

Successors to IALS added problem solving and problem solving in technology-rich environment to the assessment dimensions. Third party assessment providers also developed tests for other Essential Skills dimensions, such as writing (by Essential Skills Group), and digital literacy (by SRDC). However, these other assessments are not standardized, and have unknown psychometric properties.

The following is a discussion of the origin of the standardized literacy assessments from International Adult Literacy Survey (IALS), and the subsequent development of objective assessments of generic Essential Skills. All of the assessments discussed in this section are suitable for low-stakes assessment of initial needs and training gains at the group level. However, with the possible exception of TOWES Prime Sharp or Foundations, which claim high precision, none of the assessments is precise enough to make high-stakes individual level decisions, such as determination of job performance.

Though research using population based data has consistently shown that higher scores in generic assessments of Essential Skills such as numeracy, document use and reading are correlated with a variety of positive labour market outcomes, experimental data showing a causal link between interventions designed to raise Essential Skill levels and improved employment outcomes have only recently begun to be investigated. A key question we will continue to explore throughout the different stages of this project is what kinds of Essential Skills should be taught and measured in different training contexts; the answer will depend to a large extent on establishing links between skill gains and "downstream" employment outcomes.

Background: the origins of standardized assessments of Literacy and Essential Skills

The IALS was commissioned by Canada's National Literacy Secretariat. The Canadian IALS was designed and managed by Statistics Canada with the cooperation of various OECD countries. Educational Testing Service (ETS) was the test developer for the IALS and the methodology used expanded on ETS' earlier work on the US Young Adult Literacy Survey (YALS), Department of Labor

Literacy Survey, and National Adult Literacy Survey (NALS). Similar to previous literacy surveys in the US, the development of the IALS assessment followed a six-part framework (Kirsch 2001):

1. Defining Literacy

2. Organizing the Domain

3. Task Characteristics

4. Identifying and Operationalizing Variables

5. Validating Variables

6. Building an Interpretive Scheme.

In contrast to the old notion that individuals were either "literate" or "illiterate", IALS and subsequent literacy and essential skills measurements adopted a concept of literacy as a continuum of skills ranging from quite limited to very high. IALS defined literacy as "*the ability to understand and employ printed information in daily activities, at home, at work and in the community, to achieve one's goals and to develop one's knowledge and potential.*" (Canada, 1991)

It was significant that IALS moved away from past measurement definitions that compared one's abilities against that of the grade-level equivalence. Usually, grade-level measurement of literacy is a check against a long list of literacy topics and literacy proficiency specified under school curricula for children and youth in developmental stages. The grade-level measurement is a standard determined by experts corresponding to what a student *ought* to have to become the adult he or she *ought* to become. This view is probably consistent with a historical perspective that considered literature and literacy almost the same thing: "Literature was the book that a literate person read." (Meek, 1991). However, the conceptualization of literacy has been shifting towards more utilitarian functions, namely skills and competencies in daily activities.

The IALS definition focused on how well individuals function in society throughout their whole life and it regarded literacy as competencies that were comparable across different languages, cultures and social contexts. The IALS definition also means that measurement should focus on realistic functions in daily activities rather than explicit literacy standards. Grade-level measurements of language reading and writing abilities as well as knowledge of mathematics from tests are no longer suitable. Instead, assessments examined the performance of individuals in applying literacy skills to everyday life tasks.

The design of the IALS followed the domain organization of NALS which included prose literacy, document literacy, and quantitative literacy. Writing was considered to be similar to reading and the prose literacy scale was supposed to measure the combined domain of reading and writing. Speaking and listening were deemed too difficult to assess in IALS and they were not included.

The IALS definition of literacy with multiple domains also coincided with the application of Item Response Theory (IRT) as the psychometric framework for assessment. IRT imposes a model specifying how individuals of various literacy competencies would perform in every assessment task. Each assessment task has a level of difficulty specified along a continuum of individual competency. The probability of an individual successfully performing an assessment task decreases

as the task's level of difficulty exceeds their competency. For example, an individual is very likely to perform successfully (80% or over in IALS) for a task with a level of difficulty below the individual's competency level.

A given assessment task may also have different discriminatory powers in terms of identifying an individual with a competency above or below its level of difficulty. Because of the strong model structure, task (or item) difficulty and discriminatory parameters are independent from the sample or other assessment tasks. This is in stark contrast with traditional standard-based literacy assessment, in which scores are dependent on the assessment construction sample and tasks. The properties of IRT facilitate cross-country comparisons and usage of computerized adaptive testing techniques. More details about the psychometrics of assessments and IRT are discussed in Appendix A.

Assessment task characteristics

Besides the assessment's psychometric framework, it is important to establish the validity between assessment tasks or items and skill domains. After domain structure was determined, assessment tasks of IALS were designed and manipulated based on:

- adult contexts/content characteristics (such that no single group of adults is either advantaged or disadvantaged by the context or content included in the assessment),

- the characteristics of materials/texts (a broad range of both prose and document text types are included), and

- the characteristics of processes/strategies (the goal or purpose assessment takers are asked to assume).

These three task characteristics were operationalized as variables. In terms of context and content, tasks selected covered six categories: home and family, health and safety, community and citizenship, consumer economics, work, and leisure and recreation. For materials and texts, the design distinguished between continuous (e.g., narration) and non-continuous (e.g., matrix documents) texts. Finally, task selection considered the processes or strategies used to relate information requested in the question to the necessary information in the text, and the processes used to provide a response. The variables included type of match (through locating, cycling, integrating and generating), type of information requested and the plausibility of distracting information.

A test-taking model (the steps taken by a test taker to understand a question and respond) was applied to each task to assign values of information processing demands (from 1 to 5) to the set of variables affecting task performance in each literacy domain. The validation results suggested that task characteristics largely explained task difficulty in each domain. For prose literacy tasks, type of match was most important in predicting task difficulty, followed by plausibility of distractor and type of information. Type of match, plausibility of distractor, and document readability were most important in explaining the difficulty of document literacy tasks. The difficulty of quantitative literacy tasks was related to the task's operation specificity and plausibility of distractor.

The three domains of literacy in IALS

Prose literacy

Prose literacy is defined as the ability to understand and use information from texts such as editorials, news stories, poems and fiction. Prose tasks in IALS's assessment require respondents to first identify "given" and "requested" information. Given information is known and assumed to be true based on the way a question or directive is stated. Requested information in a question or directive is information being sought (Kirsh et al., 1998). For example, a Level 1 prose task asks the respondent to determine from a medicine label the maximum number of days one should take the medicine (Figure 1) without any distracting information. Figure 2 illustrates a Level 2 prose task, asking the reader to determine what happens when an impatiens plant is exposed to temperatures of 14 degree Celsius or lower. It is more difficult than a Level 1 task because there is another sentence in the relevant section about the general temperature requirements of the plant, providing additional information that could have distracted some readers.

Figure 1   An example of IALS Level 1 prose task

Source: p. 4 of Statistics Canada (1997).

Figure 2   An example of IALS Level 2 prose task

IMPATIENS

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

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

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

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

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

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

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

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

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

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

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

Source: p. 4 of Statistics Canada (1997).

Document literacy

Document literacy is defined as the ability to locate and use information from documents such as job applications, payroll forms, transportation schedules, maps, tables and graphs. Document literacy tends to be the principal form of literacy in non-school settings, in contrast to prose literacy (Kirsh et al., 1998). Procedural knowledge may be needed to transfer information from one source or document to another, as is necessary in completing applications or order forms. For example, Figure 3 shows one of the document tasks in IALS: a quick copy printing requisition form commonly found in the workplace. The Level 3 task associated with the document asks the reader to explain whether or not the quick copy centre would make 300 copies of a statement that is 105 pages long. The reader has to determine whether conditions stated in the question meet those provided in the guidelines to this document.

Figure 3   An example of IALS Level 3 document literacy task

Source: p. 6 of Statistics Canada (1997).

Quantitative literacy

Quantitative literacy is defined as the ability to perform arithmetic functions such as balancing a cheque book, calculating a tip, or completing an order form. IALS assessments of quantitative literacy asks respondents to perform arithmetic operations such as addition, subtraction, multiplication, or division either singly or in combination using numbers or quantities that are embedded in printed material. Although it seems the ability is fundamentally different from those involved in processing prose and documents, empirically they are related since processing printed information is essential to quantitative tasks as well (Kirsh et al., 1998). The two unique attributes of quantitative tasks: are operation specificity (the processing of identifying and entering numbers into arithmetic expression) and type of calculation.

An example of an IALS quantitative literacy assessment task asks the reader to "calculate the total amount of money you will have if you invest $100 at a rate of 6% for 10 years" using a compound interest table (Figure 4). This Level 4 task requires the reader to identify the correct amount of interest in the table and add the interest to the principal in order to obtain the correct answer.

Figure 4   An example of IALS Level 4 quantitative literacy task

**Compound Interest
Compounded Annually**

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

Source: p. 7 of Statistics Canada (1997).

Interpretations of assessment results

Task characteristics also help to interpret the assessment results. IALS scores were calibrated to a 0-500 scale with 5 levels: Level 1 (0-225), Level 2 (226-275), Level 3 (276-325), Level 4 (326-375) and Level 5 (376-500). The process variables of tasks tended to "shift" at every 50-point interval starting from 225 in North American literacy surveys (Kirsch et al., 1998). The thresholds between these five levels were chosen such that each individual's assessed level represented their ability to perform most of the tasks with corresponding information-processing demand characteristics. For example, an individual performing at 300 on a literacy scale is expected to be able to perform the average Level 1, 2 or 3 task with a high degree of proficiency. This does not mean that a Level 3 person will not be able to perform correctly on any Level 4 tasks, but rather that the person will not be able to perform Level 4 tasks with the same consistency as a Level 4 person.

Table 1   Task description for prose, document, and quantitative literacy in IALS, by level

	Prose	Document	Quantitative
Level 1 (0-225)	Most of the tasks at this level require the reader to locate one piece of information in the text that is identical or synonymous to the information given in the directive. If a plausible incorrect answer is present in the text, it tends not to be near the correct information.	Most of the tasks at this level require the reader to locate a piece of information based on a literal match. Distracting information, if present, is typically located away from the correct answer. Some tasks may direct the reader to enter personal information onto a form.	Although no quantitative tasks used in the IALS fall below the score value of 225, experience suggests that such tasks would require the reader to perform a single, relatively simple operation (usually addition) for which either the numbers are already entered onto the given document and the operation is stipulated, or the numbers are provided and the operation does not require the reader to borrow.
Level 2 (226-275)	Tasks at this level tend to require the reader to locate one or more pieces of information in the text, but several distractors may be present, or low-level inferences may be required. Tasks at this level also begin to ask readers to integrate two or more pieces of information, or to compare and contrast information.	Document tasks at this level are a bit more varied. While some still require the reader to match on a single feature, more distracting information may be present or the match may require a low-level inference. Some tasks at this level may require the reader to enter information onto a form or to cycle through information in a document.	Tasks in this level typically require readers to perform a single arithmetic operation (frequently addition or subtraction) using numbers that are easily located in the text or document. The operation to be performed may be easily inferred from the wording of the question or the format of the material (for example, a bank deposit form or an order form).
Level 3 (276-325)	Tasks at this level tend to direct readers to search texts to match information that require low-level inferences or that meet specified conditions. Sometimes the reader is required to identify several pieces of information that are located in different sentences or paragraphs rather than in a single sentence. Readers may also be asked to integrate or to compare and contrast information across paragraphs or sections of text.	Tasks at this level appear to be most varied. Some require the reader to make literal or synonymous matches, but usually the matches require the reader to take conditional information into account or to match on multiple features of information. Some tasks at this level require the reader to integrate information from one or more displays of information. Other tasks ask the reader to cycle through a document to provide multiple responses.	Tasks found in this level typically require the reader to perform a single operation. However, the operations become more varied — some multiplication and division tasks are found in this level. Sometimes two or more numbers are needed to solve the problem and the numbers are frequently embedded in more complex displays. While semantic relation terms such as "how many" or "calculate the difference" are often used, some of the tasks require the reader to make higher order inferences to determine the appropriate operation.

	Prose	Document	Quantitative
Level 4 (326-375)	These tasks require readers to perform multiple-feature matching or to provide several responses where the requested information must be identified through text-based inferences. Tasks at this level may also require the reader to integrate or contrast pieces of information, sometimes presented in relatively lengthy texts. Typically, these texts contain more distracting information and the information that is requested is more abstract.	Tasks at this level, like those in the previous levels, ask the reader to match on multiple features of information, to cycle through documents, and to integrate information; frequently however, these tasks require the reader to make higher order inferences to arrive at the correct answer. Sometimes, conditional information is present in the document, which must be taken into account by the reader.	With one exception, the tasks at this level require the reader to perform a single arithmetic operation where typically either the quantities or the operation are not easily determined. That is, for most of the tasks at this level, the question or directive does not provide a semantic relation term such as "how many" or "calculate the difference" to help the reader.
Level 5 (376-500)	Some tasks at this level require the reader to search for information in dense text that contains a number of plausible distractors. Some require readers to make high-level inferences or use specialized knowledge.	Tasks at this level require the reader to search through complex displays of information that contain multiple distractors, to make high-level inferences, process conditional information, or use specialized knowledge.	These tasks require readers to perform multiple operations sequentially, and they must disembed the features of the problem from the material provided or rely on background knowledge to determine the quantities or operations needed.

Source: p. 87 of Statistics Canada (1996).

The information-processing demand of modern workplace tasks is believed to be at Level 3. As a result, many policy studies use of the proportion of the population at Level 2 or below as a key indicator of policy or program progress. Unfortunately, the Level 2/3 threshold is sometimes misinterpreted according to the traditional binary classification of "illiterate/literate" rather than as an indicator of ability to perform most Level 3 work tasks consistently.

Properties of the IALS

One hundred and fourteen tasks were grouped into three skill domain scales and divided into seven blocks, which in turn were compiled into seven test booklets. Each booklet contained three blocks of tasks. Respondents were expected to spend 45 minutes to complete the assessment. At the beginning of the assessment, respondents were asked to perform a set of six "core" tasks, and only those who were able to perform at least two core tasks correctly proceed to the full assessment.

The application of item response theory to IALS produced not only a skill domain score for each individual but also the uncertainty associated with the score under the model. Thus instead of presenting a single individual score, IALS applied a plausible value methodology to present a range

of potential individual scores for each domain. This method aimed to produce consistent population level estimates of literacy rather than optimizing the precision of individual assessments.

As a result, there is a degree of measurement error associated with each domain score for a given individual. IALS's microdata is no longer publicly available, and the available documentation does not specify the size of measurement error. However, the successor of IALS, the Adult Literacy and Lifeskills Survey, suggested that the median standard deviation of measurement error was about 16 to 20 points, with 90 per cent of the Canadian sample having a standard deviation of measurement errors within 35 points. In other words, though many IALS participants were probably correctly assessed within one skill level, it would have been possible to classify a minority of participants into either of two adjacent levels of competency. In the worst cases, the standard deviation of measurement error went as high as 80 points, suggesting that the respondent could have been at any of the five levels. In general, IALS types of assessments are not sufficiently precise to be used for high-stake tests at the individual level (such as determining certification), and are intended more for trend analysis involving groups or sub-populations.

Literacy and Essential Skills assessments related to IALS

Adult Literacy and Lifeskills Survey (ALL)

The Adult Literacy and Lifeskills Survey (ALL) was the successor to the IALS. It retained most of the same assessment methodology for prose, document, and numeracy literacy while introducing the domain of "problem solving".

The definitions for prose and document literacy in ALL were the same as in IALS. However, numeracy's definition was broadened to "the knowledge and skills required to effectively manage and respond to the mathematical demands of diverse situations" Adult numeracy extended to a possession of number sense, estimation skills, measurement and statistical literacy. It included proficiencies in not just commonly encountered situations but also new situations, and required a person to be able to communicate mathematical information and reasoning processes.

For the development of numeracy assessment items, a broader definition of numerate behaviour was developed to serve as the basis: "Numerate behaviour is observed when people manage a situation or solve a problem in a real context; it involves responding to information about mathematical ideas that may be represented in a range of ways; it requires the activation of a range of enabling knowledge, factors and process." (Statistics Canada, 2011). As a result, assessment tasks were not restricted to numbers embedded in printed materials, but also included a wider range of situations that presented respondents with mathematical information. For example, Figure 5 illustrates a task that asked respondents to describe how the amount of Dioxin (a toxin found in fish in Baltic Sea) changed from 1975 to 1995. The respondents were not required to calculate the amount of change over each of the periods, just describe in their own words the change in levels.

Figure 5   An example of ALL numeracy task

Is breast milk safe?

Since the 1970s, scientists have been worried about the amount of Dioxin, a toxin in fish caught in the Baltic sea. Dioxin tends to accumulate in breast milk and can harm newborn babies.

The diagram shows the amount of Dioxin in the breast milk of North European women, as found in studies done from 1975 to 1995.

Amount of Dioxin in Breast Milk
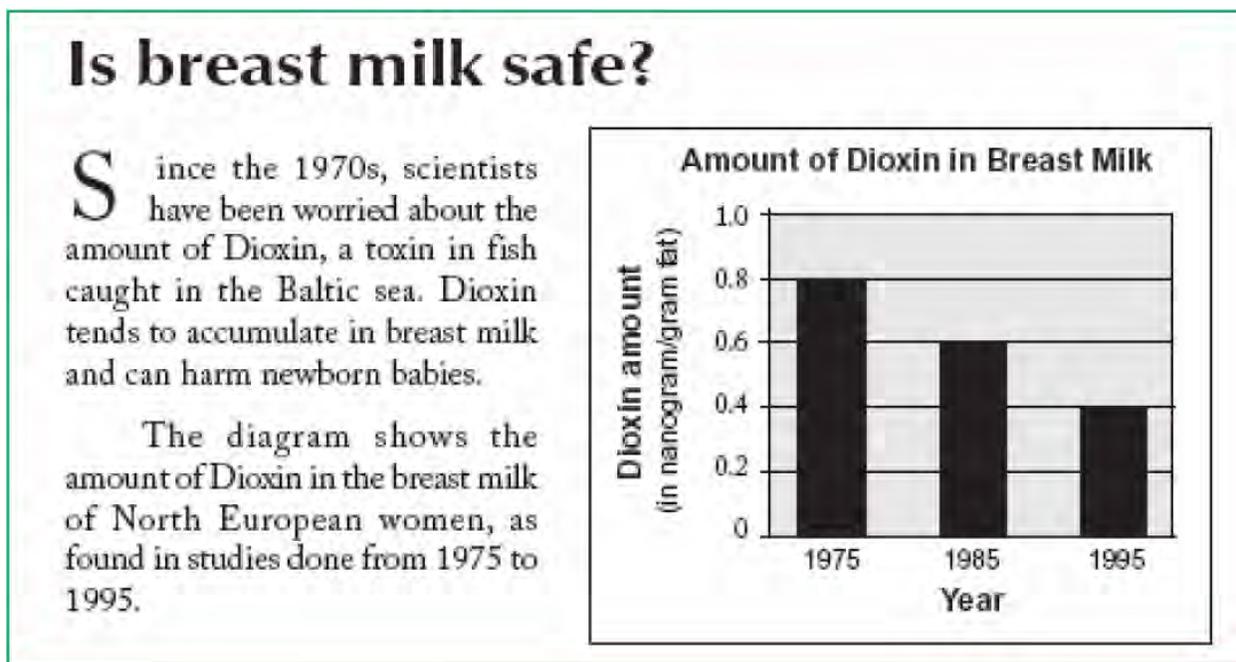
Source: p. 28 of Statistics Canada (2011).

ALL also assessed a new literacy domain not found in the IALS: problem solving. Statistics Canada (2011) defines problem solving as:

> *"Problem solving is goal-directed thinking and action in situations for which no routine solution procedure is available. The problem solver has a more or less well-defined goal, but does not immediately know how to reach it. The incongruence of goals and admissible operators constitutes a problem. The understanding of the problem situation and its step-by-step transformation, based on planning and reasoning, constitute the process of problem solving."*

ALL focused on a subset of problem solving: analytical problem solving, which involves the following steps: 1. Searching for information, and structuring and integrating it into a mental representation of the problem ("situational model"), 2. Reasoning, based on the situational model, and 3. Planning actions and other solution steps.

The pool of problem solving tasks was designed to assess individual's performance in each of the five stages of problem solving: define the goal; analyze the given situation and construct a mental representation; devise a strategy and plan the steps to be taken; execute the plan; and evaluate the result.

In ALL, four levels of problem-solving proficiency are postulated:

Table 2   Task description for problem solving skills in ALL, by level

	Problem solving
Level 1 (0-225)	At a very elementary level, concrete, limited tasks can be mastered by applying content-related, practical reasoning. At this level, people will use specific content related schemata to solve problems.
Level 2 (226-275)	The second level requires at least rudimentary systematical reasoning. Problems at this level are characterized by well-defined, one-dimensional goals; they ask for the evaluation of certain alternatives with regard to transparent, explicitly stated constraints. At this level, people use concrete logical operations.
Level 3 (276-325)	At the third level of problem-solving proficiency, people will be able to use formal operations (e.g., ordering) to integrate multi-dimensional or ill-defined goals, and to cope with non-transparent or multiple dependent constraints.
Level 4	At the final and highest level of competency, people are capable of grasping a system of problem states and possible solutions as a whole. Thus, the consistency of certain criteria, the dependency among multiple **sequences of actions and other "meta-features" of a problem situation may be considered systematically. Also,** at this stage people are able to explain how and why they arrived at a certain solution. This level of problem-solving competency requires a kind of critical thinking and a certain amount of meta-cognition.

Assessment of analytical problem solving in ALL was implemented by using a project approach. The following is an example of a project from the ALL's Users Guide.

Imagine that you live in City A. Your relatives are scattered throughout the country and you would like to organize a family reunion. The reunion will last 1 day. You decide to meet in City B, which is centrally located and accessible to all. Since you and your relatives love hiking, you decide to plan a long hike in a state park close to City B. You have agreed to be responsible for most of the organization.

Example task: Set the date for the reunion

The family reunion should take place sometime in July.

You asked all your relatives to tell you which dates would be suitable. After talking to them, you made a list of your relatives' appointments during the month of July. Your own appointment calendar is lying in front of you. You realize that some of your relatives will have to arrive a day early in order to attend the family reunion and will also only be able to return home on the day after the meeting.

Please look at the list of your relatives' appointments and your own appointment calendar.

List of your relatives' appointments in July 1999

Henry	Karen	Peter	Janet	Anne	Frank
Vacation in City E beginning July 16; Appointment on July 11	Every day of the week is okay except Thursdays and on July 16	Business appointments on July 2, July 13, and between July 27 and 29	Doesn't have any appointments	Unable to attend reunion on July 5, July 20, or July 24	Has be to away sometime during the 1 full week in July on business, but will find out the exact dates shortly before

Henry, Karen, and Peter could arrive on the same day as the reunion whereas Janet, Anne, and Frank can only arrive on the afternoon before and return home on the day after the reunion.

Your appointment calendar for July 1999

Sun	Mon	Tue	Wed	Thu	Fri	Sat
27	28	29	30	Jul 1 — Meeting with David	2	3
4	5	6	7	8	9	10 — Hike in City C
11	12	13	14	15	16	17
18	19	20	21	22	23	24
25	26	27	28 — Vacation	29 — Vacation	30 — Vacation	31

Question 1. Which of the following dates are possible for the family reunion?

Please select all possible dates.

a. July 4

b. July 7

c. July 14

d. July 18

e. July 25

f. July 29

The ALL assessment included four 30-minute blocks of items (for prose and document literacy), two 30-minute blocks of numeracy items, and two 30-minute blocks of problem-solving items. The blocks of assessment items were organized into 28 task booklets according to a balanced in block design.

Prose, document, numeracy and problem solving were all scored on 0-500 scales. The levels and interpretations of prose and document literacy were the same as they had been for the IALS, while ALL's numeracy was consistent with but expanded upon IALS' quantitative literacy interpretation. In psychometric terms, ALL's assessments were created using the 2 parameter logistic IRT and each domain score was presented as 5 plausible values for each individual. The following is the table of the measurement imprecision as measured by the minimum, 10th, 50th, 90th percentile, and maximum of the individual standard deviation of the plausible values.

Table 3   Summary of individual standard deviation of plausible values in ALL assessments

Domain	Minimum	10th percentile	50th percentile	90th percentile	Maximum
Prose	0.9	8.4	16.4	31.1	75.2
Document	1.2	8.8	17.2	32.1	76.6
Numeracy	1.1	10.4	19.7	34.5	80.1
Problem-solving	0.5	10.0	18.6	30.4	68.9

Note: Calculation by SRDC using ALL 2003 PUMF.

Similar to IALS, the ALL prose and document literacy tasks were later scored on a combined domain called literacy. Also similar to IALS, Educational Testing Services was involved in the development of the assessment tools used in ALL.

Re-using some of the assessment questions from the IALS or ALL is possible, but only by commissioning Statistics Canada to conduct a customized survey. The pricing, however, is not clear. In addition, the typical turn around time from survey design to data collection and post-collection processing is over one year. Therefore, these kinds of measures would only be useful for population-based research, rather than for example monitoring group starting points and skill gains during training.

Programme for the International Assessment of Adult Competencies (PIAAC)

The Programme for the International Assessment of Adult Competencies (PIAAC) was the successor to ALL. It retained mostly the same assessment methodology for literacy (combining both prose and document), and numeracy. The definitions for prose and document literacy and numeracy in PIAAC were the same as in ALL, and the assessments for these domains were linked to those of ALL to facilitate comparison.

Instead of analytical problem solving, PIAAC assessed respondents' problem solving in technology-rich environments (PSTRE), a domain related to digital essential skills. Problem solving in technology-rich environments is defined as "using digital technology, communication tools and networks to acquire and evaluate information, communicate with others and perform practical tasks." (OECD, 2013)

In its implementation, the PSTRE domain focused on "abilities to solve problems for personal, work and civic purposes by setting up appropriate goals and plans, and accessing and making use of information through computers and computer networks". Specifically, the PSTRE covered the class of problems people deal with when using information and computer technology (ICT). The problems share three characteristics:

- The existence of the problem is primarily a consequence of the availability of new technologies

- The solution to the problem requires the use of computer-based artifacts

- The problems are related to the handling and maintenance of technology-rich environments themselves.

The core dimensions of PSTRE included task/problem statements (the circumstances that trigger a person's awareness and understanding of the problem and determine the actions needed to be taken in order to solve the problem), technologies (devices, applications and functionalities through which problem is solved), and cognitive dimensions (the mental structures and processes involved when a person solves a problem). Because PSTRE aimed to encompass more than purely instrumental skills (the so-called information, computer and technology (ICT) skills) related to the knowledge and use of digital technologies, the cognitive dimensions of problem solving were considered the central object of the assessment with the use of ICT as secondary.

Tables 4 and 5 present the task descriptions in PIAAC by domain and level.

Table 4   Task description for literacy and numeracy in PIAAC by level

	Literacy (prose and document combined)	Numeracy
Below level 1 (0-175)	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 non-continuous. Tasks below Level 1 do not make use of any features specific to digital texts.	Tasks at this level are set in concrete, familiar contexts where the mathematical content is explicit with little or no text or distractors and that require only simple processes such as counting, sorting, performing basic arithmetic operations with whole numbers or money, or recognizing common spatial representations.
Level 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 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 non-continuous 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.	Tasks in this level require the respondent to carry out basic mathematical processes in common, concrete contexts where the mathematical content is explicit with little text and minimal distractors. Tasks usually require simple one-step or two-step processes involving, for example, performing basic arithmetic operations; understanding simple percents such as 50%; or locating, identifying and using elements of simple or common graphical or spatial representations.
Level 2 (226-275)	At this level, the complexity of text increases. The medium of texts may be digital or printed, and texts may comprise continuous, non-continuous 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.	Tasks in this level require the respondent to identify and act upon mathematical information and ideas embedded in a range of common contexts where the mathematical content is fairly explicit or visual with relatively few distractors. Tasks tend to require the application of two or more steps or processes involving, for example, calculation with whole numbers and common decimals, percents and fractions; simple measurement and spatial representation; estimation; and interpretation of relatively simple data and statistics in texts, tables and graphs.

	Literacy (prose and document combined)	Numeracy
Level 3 (276-325)	Texts at this level are often dense or lengthy, including continuous, non-continuous, 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.	Tasks in this level require the respondent to understand mathematical information which may be less explicit, embedded in contexts that are not always familiar, and represented in more complex ways. Tasks require several steps and may involve the choice of problem-solving strategies and relevant processes. Tasks tend to require the application of, for example, number sense and spatial sense; recognizing and working with mathematical relationships, patterns, and proportions expressed in verbal or numerical form; and interpretation and basic analysis of data and statistics in texts, tables and graphs.
Level 4 (326-375)	Tasks at this level often require respondents to perform multiple-step operations to integrate, interpret, or synthesize information from complex or lengthy continuous, non-continuous, mixed, or multiple type texts. Complex inferences and application of background knowledge may be needed to perform successfully. Many tasks require identifying and understanding one or more specific, non-central 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.	Tasks in this level require the respondent to understand a broad range of mathematical information that may be complex, abstract or embedded in unfamiliar contexts. These tasks involve undertaking multiple steps and choosing relevant problem-solving strategies and processes. Tasks tend to require analysis and more complex reasoning about, for example, quantities and data; statistics and chance; spatial relationships; change; proportions; and formulas. Tasks in this level may also require comprehending arguments or communicating well-reasoned explanations for answers or choices.
Level 5 (376-500)	At this level, tasks may require the respondent to search for and integrate information across multiple, dense texts; construct syntheses of similar and contrasting ideas or points of view; or evaluate evidence-based arguments. Application and evaluation of logical and conceptual models of ideas may be required to accomplish tasks. Evaluating reliability of evidentiary sources and selecting key information is frequently a key requirement. Tasks often require respondents to be aware of subtle, rhetorical cues and to make high-level inferences or use specialized background knowledge.	Tasks in this level require the respondent to understand complex representations and abstract and formal mathematical and statistical ideas, possibly embedded in complex texts. Respondents may have to integrate multiple types of mathematical information where considerable translation or interpretation is required; draw inferences; develop or work with mathematical arguments or models; and justify, evaluate and critically reflect upon solutions or choices.

Source: OECD (2013).

Table 5   Task description for Problem solving in technology-rich environments in PIAAC by level

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

Source: OECD (2013).

PIAAC was also the first international survey on literacy and skills that made use of computerized adaptive testing (i.e., streaming test takers to level-appropriate questions based on their performance on previous questions) to improve efficiency. Because respondents required a working level of computer usage skills to complete the full assessment using a computer, respondents with no or extremely limited experience with the use of computers were given a paper-and-pencil based instrument (with only the literacy and numeracy assessments). Otherwise, respondents would respond to the survey through the computer-based instrument (including PSTRE), under the supervision of the interviewer.

PIAAC assessments were designed based on an assumption of 60 minutes of testing time on average. However, it was not a timed assessment and some respondents were expected to take longer to complete the survey. The paper-delivered branch of assessment included a 10-minute core assessment of literacy and numeracy skills. Those who performed at or above a threshold were randomly assigned to a 30-minute cluster of literacy or numeracy items, followed by a 20-minute assessment of component skills. Those below the threshold were assessed for the reading component skills measure.

The computer delivered branch of assessment was similar. Respondents were first assessed for a computer based assessment core section composed of two 5-minute stages. Poor performance on either stage resulted in switching over to the paper-and-pencil instruments. Those with adequate performance at both stages of the core section proceeded to the full computer based assessment. Levels of questions asked were determined based on each respondent's assessed proficiency at the first stage of the computer based assessment (adaptive testing).

In terms of psychometrics, PIAAC's assessments were also created using item response theory and each domain score was presented as 10 plausible values for each individual. The following is a table of the measurement imprecision as measured by the minimum, 10th, 50th, 90th percentile, and maximum of the individual standard deviation of plausible values. The majority of respondents' assessments were precise within a level.

Table 6   Summary of individual standard deviation of plausible values in PIAAC assessments

Domain	Minimum	10th percentile	50th percentile	90th percentile	Maximum
Literacy	2.5	9.8	15.4	24.8	60.1
Numeracy	3.9	10.8	17.1	28.0	66.4
Problem-solving in technology reach environment	3.5	10.6	16.3	23.8	47.5

Note: Calculation by SRDC using PIAAC microdata.

The full set of PIAAC assessments are available from the OECD's Education and Skills Online (ESO): http://www.oecd.org/skills/ESonline-assessment/abouteducationskillsonline/. This is the tool being used by the Higher Education Quality Council of Ontario's Essential Adult Skills Initiative. It costs €5 to €11 per assessment depending on the number of assessments and the assessment package.

If a customized assessment is needed, it is possible to commission Statistics Canada to conduct a customized survey that includes assessment questions from PIAAC. The pricing, however, is not clear, and again the turn around time from survey design to data collection and post-collection processing is over one year, limiting the utility of these kinds of customized surveys.

It might be easier to commission Educational Testing Services (ETS) to conduct customized assessments, since ETS was involved in the development of the assessment tools used in PIAAC as well as IALS and ALL.

Educational Testing Services (ETS) tests and assessments

With their expertise in developing assessments of "core" literacy, ETS has been providing literacy and skills assessment products for general usage. There was a web based literacy assessment called Canadian Literacy Evaluation (CLE) up to the year 2010 that evaluated an individual's literacy skill based on the IALS scale. CLE was developed with support from the Learning Policy Directorate of then HRSDC to be a Canadian IALS/IALSS-based self-assessment test in both English and French. The psychometric properties of the assessment were not disclosed, though one could reasonably assume that ETS applied the same methodologies and similar test items as they had used in developing IALS and ALL.

CLE consisted of a set of background questions and three sets of literacy tasks: prose, document and quantitative. The test took about 90 minutes to complete, however there was no time limit. Its scores described an individual's strengths and weaknesses with respect to the types of literacy tasks they could perform. The scores could be used to characterize an individual's performance in terms of national and international results. However, CLE was not intended to test course materials or learning of curriculum, and was not very informative for adults who had difficulty reading.

Partner provincial organizations who wished to collaborate with HRSDC on the CLE would sign a research agreement with the Learning Policy Directorate and agree to follow the research protocol so that the data would be reliable enough to compare across programs, policies and people. The research agreement (Memorandum of Understanding) allowed organizations and/or provinces to link their project to their policy concerns and to customize it to their needs: it could focus on a sector (e.g.: construction), a group (e.g., unemployed, immigrants), or a process of their choice. There was no restriction regarding the provider of the intervention. The tests and analysis were made available under the agreement, however, all costs of training and participation were assumed by the partner organization.

Figures 6 and 7 illustrate screen captures from the CLE. Internet Explorer was required to access the platform at the time. However, the CLE platform was not updated to match with the latest browser technologies in 2010 and it was subsequently discontinued because of incompatibility. It is unclear to what extent the content of the CLE was used for ESDC's current Essential Skills indicator, which includes online assessments of reading, document use, and numeracy (for a more detailed description of the Essential Skills indicator, see p. 35).

Currently, ETS in the US has an online assessment product called *WorkFORCE*® Assessment for Cognitive Ability that assesses individual's proficiencies in Prose, Document and Quantitative literacy. It was developed based on ETS's work and items on literacy assessments of IALS and its successors.

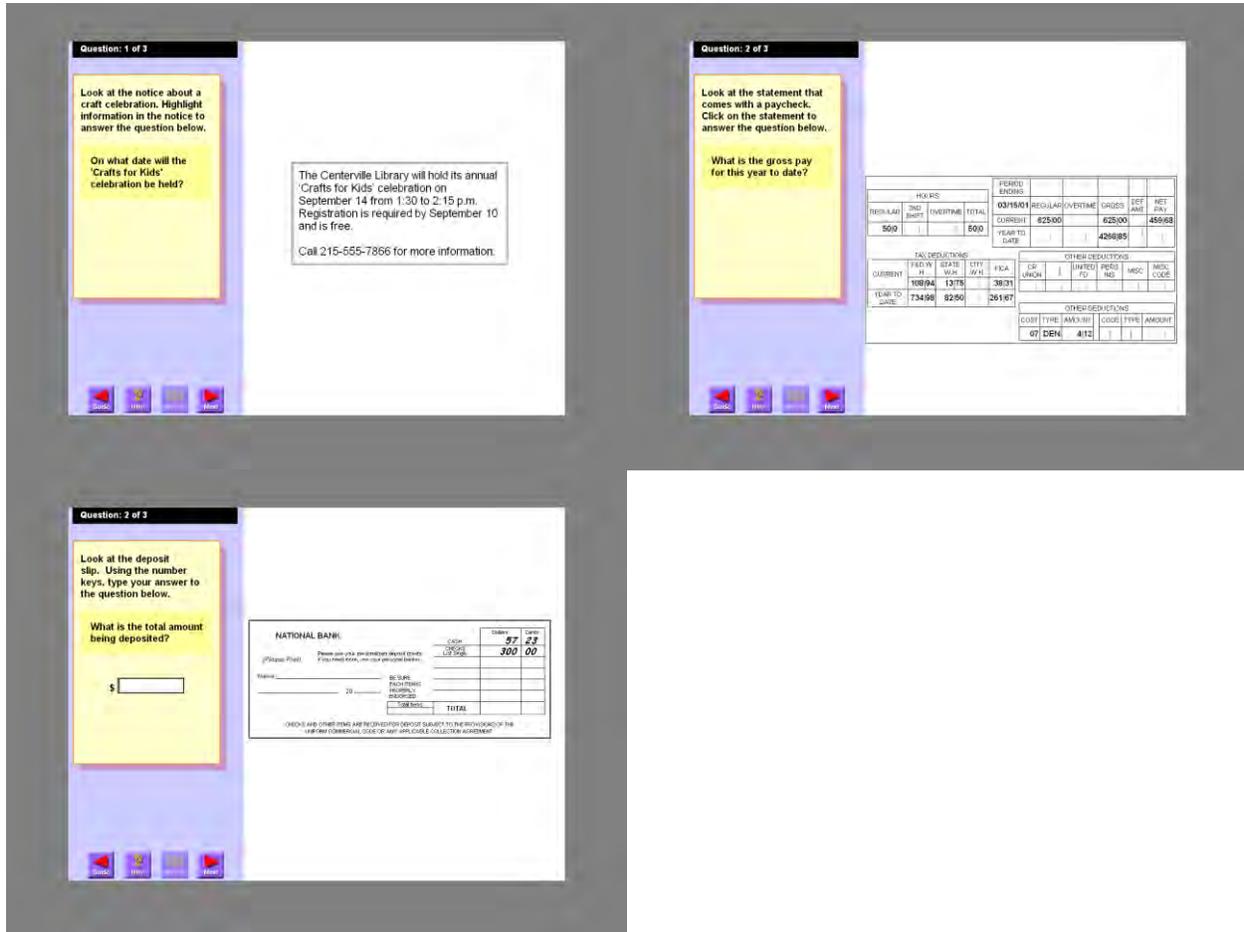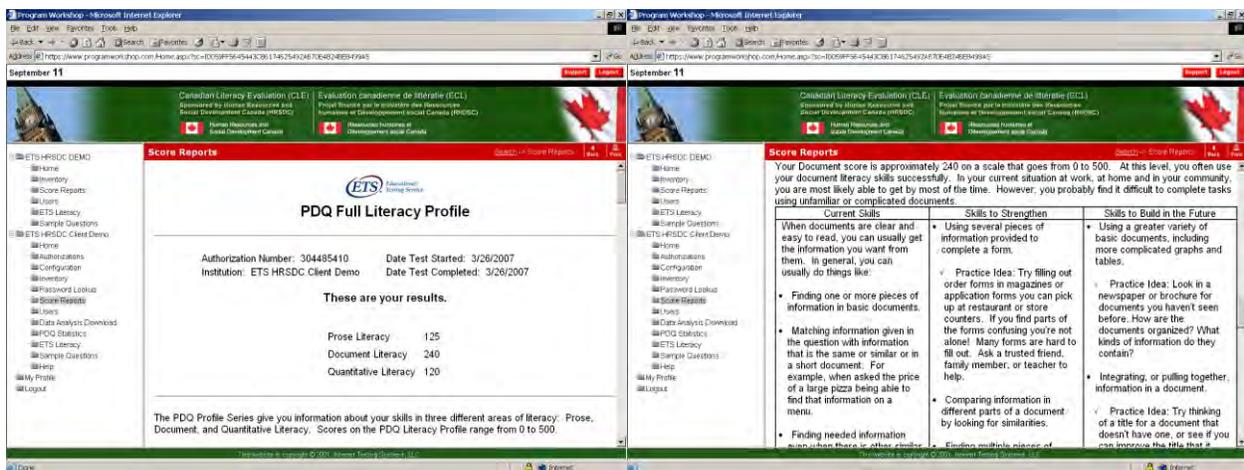Figure 6   Examples of CLE assessment items on prose, document, and quantitative literacy

Figure 7   Examples of CLE score report

Bow Valley College's Test of Workplace Essential Skills (TOWES)

TOWES measures essential or employability skills in three domains – reading text, document use and numeracy. TOWES developers made use of the data collected during HRSDC's Essential Skills Research Project (ESRP) on actual tasks carried out by workers in Canada. The development shared some similarity with that of IALS, and test results are delivered in a framework based on IALS levels (even though domain definitions are closer to ALL than IALS) and are consistent from one occupation to another. Figure 8 presents one of the questions from an example test booklet provided by Bow Valley College.

Figure 8   A sample question from TOWES

TOWES developers stated that TOWES assessments adhere to internationally established psychometric guidelines to guarantee the quality of the assessment. The test items in their database are kept confidential. Yamamoto and Kirsch (2002) showed that of the 412 questions examined (305 TOWES and 107 IALS) at the time, those from TOWES were linked to those from IALS assessments. The linkage and validation study showed that TOWES questions exhibited a very high proportion of agreement in rescoring (0.97), non-significant effects of item order, and very similar characteristics between TOWES and IALS items.

Because TOWES was developed based on item response theory, each individual assessment result can be accompanied by a standard error of measurement. The precision of paper-based assessment depends on the customization required. In general, the longer the assessment, the more precise the measurement and the smaller the standard error. Also, test booklets provided by TOWES target specific skill levels. For example, if most of the test takers are expected to be at Level 2, TOWES booklets will select more Level 2 questions or questions closer to Level 2. A Level 2 booklet would be less precise in assessing individuals at Level 3 or Level 1, compared to booklets corresponding to their respective skill levels.

The web-based version of TOWES (TOWES Prime) uses adaptive testing to make sure that adequate measurement precision is achieved (illustrated in Figure 9, though precision is mislabeled as "accuracy"). There are 4 levels of precision: Foundation has a very high level of precision with a +/- 5 points expected error in measurement, with Sharp next at +/- 10 points, and Focus at +/- 15 points. The Locators is the least precise version with a +/- 25 points expected error in measurement (which is comparable to PIAAC/ALL).

Figure 9   Precisions and typical usages of TOWES Prime products

Product	Accuracy (expected error)	Contexts represented in the Assessment	Typical Usage
Foundation	High (+/- 5)	Cognitive domains associated with processing linguistic information	Diagnose reading problems or issues that affect the ability of an individual to learn how to read.
Focus (Locators)	Low (+/- 25)	Representative of everyday life and the workplace, and may be customised for specific contexts (e.g., oil field, service industries, healthcare, education)	Low-stakes testing where the results may be used with other information to inform decisions that affect individual outcomes or where the results are interpreted in aggregate form (i.e., statistics) to inform decision-making about large groups of people.
Focus	Medium (+/- 15)		
Sharp	High (+/- 10)		High stakes testing where the results are the primary source of information used to make decisions that affect individual outcomes.

The developers' stated purpose of TOWES Prime is to "ensure that assessment experiences deliver accurate and valid results, a necessary component for using test data to make pro-active decisions" (Cartwright, TOWES & Murray, 2014).

In addition to scores in assessed literacy domains, participants are also provided with *Session Reports*, including:

- the location of their assessment;
- the duration;

- the probability of the participant attaining each of the OECD proficiency levels;

- the margin of error for the IALS score; and

- the participant's percentage of correct questions, including a 95% confidence interval.

The document also provides a brief overview of the scales, a longer description of the OECD level achieved, and some suggestions for improvement, including links to a *Learning Recommendations* document that is general to each OECD level, including some additional instructional recommendations.

Administrative reports accessible on the TOWES online portal include the date and location of assessment, duration, OECD proficiency level, OECD placement confidence, IALS scale score and standard error, and percentage of correct answers, with a 95% confidence interval.

Training for TOWES PRIME test administration – referred to as invigilation – is mandatory (this is similar to TOWES paper-version's administration); it is delivered either in person or via conference call and online web-meeting. Invigilators administering TOWES PRIME products must participate in an *Introduction to TOWES PRIME* session*, and *Invigilator Certification* in order to become certified to invigilate. During the testing process, each invigilator may supervise no more than 20 clients at a time, and the recommended group size is 15 clients or less. Assessment sessions with more than 20 clients must use additional invigilators, maintaining the ratio of 1 invigilator to a maximum of 20 clients.

Following completion of a consent form, invigilators lead learners through a tutorial and practice questions for TOWES PRIME as the first step of the assessment process. Completion of the practice questions is considered a necessary indicator of sufficient technical and literacy skills to continue with the assessment. In the event that participants are unable to complete the practice test, TOWES recommends that participants be issued a paper-based assessment as an alternative.

TOWES PRIME is delivered through an on-line platform that requires computers to run the web browser Google Chrome in kiosk mode. Kiosk mode prevents users from accessing the Internet for anything other than the assessment and blocks system functions that utilize CTRL+ keys and right-click menus, such as copying and zooming. This technical choice aims to prevent those with advanced computer skills from achieving a benefit unrelated to their LES levels, and to avoid confusion among test takers with low computer skills. The TOWES PRIME user interface also makes use of an alternative design for scrolling with the intention of facilitating use by participants with low computer skills.

TOWES PRIME questions are not in multiple choice format, but rather require learners to make use of different answering tools including, for instance, a highlighting tool to identify the correct response, a select tool that allows selecting the right answer, and a keyboard tool for typing in numerical or text answers.

The TOWES PRIME assessment is fully adaptive; that is, it streams test takers to level-appropriate questions based on their performance on previous questions. Answers can be reviewed and altered, although changes in answers will not alter the adaptive programming that has taken place up until

that point. Unanswered questions are marked as incorrect. Participants are able to use paper, writing tools, and calculators as needed.

In 2015, SRDC conducted an assessment of the feasibility of using TOWES Prime Focus in the context of a specific project. A non-trivial proportion of participants in a small sample SRDC recruited from a local community college faced technical difficulties when attempting to use TOWES Prime. The majority of participants with greater than high school education were assessed with a Level 1 proficiency in document use (lower than expected). Some immigrants also performed poorly in TOWES Prime relative to the other assessment given. In addition, assessment results unexpectedly failed to predict self-reported literacy practices. TOWES Prime's accuracy might have been improved since SRDC's assessment, though potential users should revisit the properties of the web-based assessment to understand where it would be applicable.

Essential Skills Group's Essential Skills Assessments and Training

The Essential Skills Group (ESG) has developed assessments for reading, document use, numeracy, oral communication, and writing. Similar to IALS, the assessments were developed based on item response theory, though it was based on a simpler Rasch/one parameter logistic model. A confidential document provided by ESG outlined a plan to extend the psychometric framework to a three parameter logistic model. ESG stated that validation had been an ongoing process, and that three third-party psychometric reviews had been completed for reading, document use and numeracy, though there are no publicly available documents about the assessments' validation. Specifically, patterns of assessment results from the ESG tool correspond very closely to those established in the 2003 Adult Literacy and Lifeskills Survey (ALL). This correspondence suggests that the ESG tool possesses a high level of external validity.

ESG has a test of writing proficiency (using a single exercise) that is suitable for low-stakes testing where the results are used for information purposes only. However, the assessment items have not been validated and the psychometric properties of the writing assessment are not known. The following description of ESG's products focus on their validated assessments in reading, document use, and numeracy.

Instructions for administering ESG assessments of the three core essential skills are provided in the "ESG Test Administrator Guide". Participants are encouraged to complete a practice assessment prior to the testing session. ESG assessments are delivered through an on-line platform that is compatible with a number of common web browsers, including Internet Explorer, Chrome, Safari, Firefox, and Opera.

The ESG assessment is composed of multiple choice questions (See Figure 10 for an example). It is adaptive; that is, it streams test takers to level-appropriate questions based on their performance in the previous section. With that in mind, answers can be reviewed and altered within each section, up until the participant completes that section. Unanswered questions are marked as incorrect. Participants are able to use paper, writing tools, and calculators as needed. The role of the administrator is simply to ensure that learners are accessing the online resources, and answering any questions about process. In practice, the ESG can be done remotely, without presence of an administrator, if appropriate for the circumstance.

Short versions of the assessment include as few as 15 questions for either reading or document use literacy and 24 questions for numeracy. Longer versions are composed of 30 questions for either reading or document use literacy and 72 items for numeracy. The assessments were designed for lower stakes testing where the results may be used with other information to inform decisions that affect individual outcomes (e.g., need for training, determining mastery) and/or where the results are interpreted in aggregate form to inform decision-making about large groups of people. The assessment is not designed to test the skills of particular sub-populations, although work is reported to be underway to develop tools for test takers with lower skill levels.

Figure 10   **An example of ESG's assessment question**
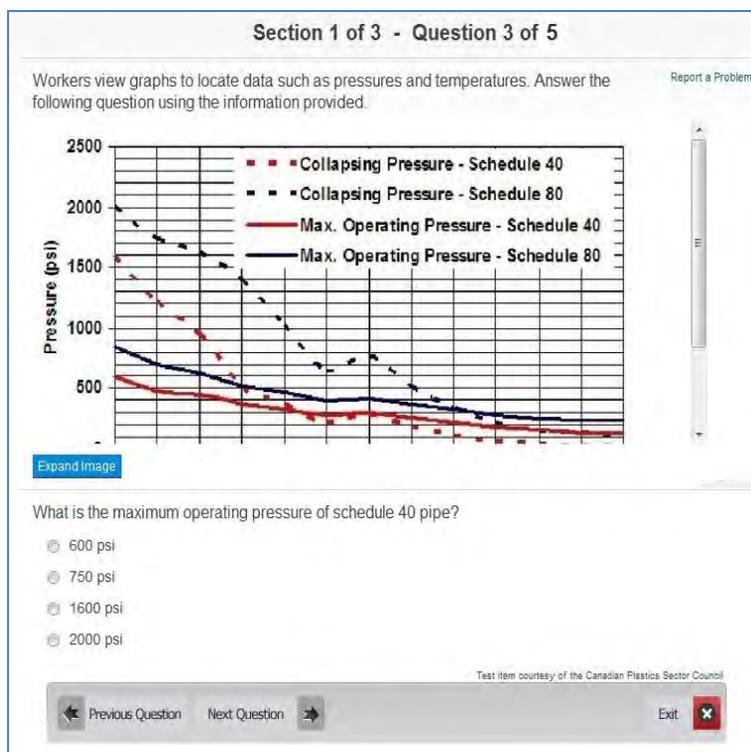
In general, ESG's assessments are geared towards respondents at proficiency Levels 1 and 2. The standard errors of measurement for a level 2 individual using shorter tests were +/- 29/29/24 points for reading, document use, and numeracy respectively. Measurement errors are lower when longer versions of each assessment are used.

In 2015, SRDC conducted an assessment of the feasibility of using ESG in the context of a specific project. In general, ESG's assessment results exhibit the expected patterns and validity. Based on SRDC's experience, ESG's assessments tend to show a higher score than TOWES' at the lower end, possibly reflecting the ESG's focus on measuring the performance of less proficient respondents.

Digital Essential Skills assessment unrelated to IALS

Given the lack of assessment instruments to measure information, computer, and technology literacy, SRDC has partnered with organizations like SkillPlan and the Restigouche Canada Business Development Corporation (CBDC) to develop items related to the assessment of various digital skills. For example, in the Workplace Digital Essential Skills (DES) Pilot Project, CDBC created an online digital skills training platform where learners who indicated no need to improve their self-assessed digital capacities were directed to an instrument where their knowledge and skills could be corroborated through objective assessment based on performance in contextualized learning situations. In parallel, SRDC developed an assessment instrument with items for each of eight generic digital tasks covering a range of technical and cognitive digital activities typically performed in the workplace. These are presented in Table 7, grouped into three main categories: Use, Understand and Create/Communicate. These tasks were identified based on available research and documentation, and validated through organizational needs assessments (ONA) with employers participating in the project. The ONA's also informed training curricula to be developed for and contextualized to occupations in four areas where research indicated that workers in rural small businesses needed basic digital skills in their jobs, namely administrative occupations, client services occupations, production occupations, and operations occupations.

Table 7   Digital Essential Skills/Tasks

No.	Description	Number of questions
	USE	
1	Refer to documents on the Internet or by using other digital media	4
2	Use documents and databases on the Internet or by using other digital media	2
3	Complete forms on the Internet or by using other digital media	3
	UNDERSTAND	
4	Perform online commercial transactions	2
5	Search databases on the Internet or other digital media to find useful information	3
6	Seek, find and choose information on the Internet	2
	CREATE/COMMUNICATE	
7	Use a calendar on the web or other digital media to schedule activities	2
8	Communicate electronically with co-workers, suppliers and clients (to coordinate workplace activities, etc.)	6

Each assessment instrument contains a series of questions of increasing complexity, where complexity is measured using expert-defined concepts of "open vs. closed" and "well vs. ill-defined" approaches: (1) closed problems are those where the resources and solutions are limited while the open ones have unlimited resources and solutions; and (2) well-defined problems are directive and clearly tell the employee what they are to do, while ill-defined problems are less clear about what to do. These are illustrated in Figure 11 below, with complexity moving in two directions: horizontally from closed to open (1 to 3) and vertically from well-defined (level 1, Use) to ill-defined (level 3,

Create). Combined, they indicate complexity levels from 1 to 5, shown in the yellow highlighted boxes. Amplifying these complexity scales are further dimensions of digital complexity shown in Figure 12 with descriptions of these dimensions shown in the first column.

Figure 11   Complexity Scale Matrix **of SRDC'S DES** assessment

(Yellow-highlighted number in boxes indicates Complexity Score)

CLOSED → OPEN

WELL DEFINED → ILL DEFINED

		1	2	3
1 USE		*Use of Digital Information for consultation purposes only* through simple instructions, repetitive action using a few simple identified technological applications within a precise task setting **[1]**	*Use of Digital Information to perform a simple task* through more complex instructions but fixed actions using several identified simple technological applications within a defined task setting **[2]**	*Use of Digital Information to perform multiple, undetermined and complex tasks* through several complex instructions or absence of instructions and unidentified actions using unlimited and undetermined complex technological applications (selection) within a familiar task setting **[3]**
		TASK: REFER TO DOCUMENTS ON THE INTERNET OR USING OTHER DIGITAL MEDIA (I.E. LETTERS, NOTES, MANUALS, ETC.)	TASK: USE DOCUMENTS/DATABASES ON INTERNET OR OTHER DIGITAL MEDIA (I.E. ADDRESS BOOKS, INVENTORIES, SUPPLIERS LISTS, ETC.)	TASK: COMPLETE FORMS ON THE INTERNET OR OTHER DIGITAL MEDIA (I.E. INVOICES, PURCHASE ORDERS, SHIPPING SLIPS, ETC.)
2 UNDERSTAND		*Understand and Organize Digital Information for one specific and simple purpose* through clear instructions and requirements using a few simple well-known technological applications within a less precise task setting **[2]**	*Understand and Organize Digital Information for simple tasks* through more complex instructions and requirements using several potential technological applications (selection) within a partially defined task setting **[3]**	*Understand and Organize Digital Information to perform multiple, undetermined and complex tasks* through several complex instructions or absence of instructions and requirements using unlimited and undetermined technological application (selection) within an less determined task setting **[4]**
		TASK: PERFORM ONLINE COMMERCIAL TRANSACTIONS (I.E. ORDER, PAY)	TASK: SEARCH DATABASES ON THE INTERNET OR OTHER DIGITAL MEDIA AND FIND USEFUL INFORMATION (I.E. SUPPLIERS LISTS, QUANTITIES, ETC.)	TASK: SEEK, FIND AND CHOOSE INFORMATION ON THE INTERNET (I.E. PRICES, GOVERNMENT POLICIES, REGULATIONS, EQUIPMENT REPAIR SPECIFICATIONS, ETC.)
3 CREATE AND DO		*Create Digital Information for one set or fixed outcomes* through simple instructions and a simple process using a few simple technological applications (selection) within a imprecise task setting **[3]**	*Create and Analyse Digital Information for several expected outcomes* through more complex instructions and multiple yet simple processes (selection) using multiple identified potential technological applications (selection) within undefined task setting **[4]**	*Create, analyse and communicate Digital Information for several self-selected potential outcomes* through several complex instructions or absence of instructions using multiple undetermined and unlimited processes and technological applications within an open undefined task setting (selection) **[5]**
		TASK: USE A CALENDAR ON THE INTERNET OR OTHER DIGITAL MEDIA (I.E. ACTIVITIES, MEETINGS, ORDERS, ETC.)	TASK: COMMUNICATE ELECTRONICALLY WITH CO-WORKERS, SUPPLIERS AND CLIENTS TO COORDINATE WORKPLACE ACTIVITIES	

Figure 12 **Dimensions of Digital Complexity Matrix for SRDC's DES** assessment

Digital Task Complexity Dimension		1 (low)	2 (medium)	3 (high)
Prescription of execution:	the clarity of the instructions provided and the number of steps to perform the task	• Simple and Precise Instructions • Repetitive actions	• The task is simple but the instructions include more steps (3 to 6 steps) • The actions are determined.	• The task and the instructions are complex and include several steps (6 + steps)/ or absence of instructions • The actions are undetermined.
Expected Outcome(s):	the clarity of the desired outcome and the impact of the outcome	• The outcome is simple and clearly defined. • *The outcome has a small impact on the workplace.*	• The outcome is less precise (expected and not defined). • *The outcome has a significant impact on the workplace.*	• The outcome is imprecise (self-selected or potential). • *The outcome has a lot of impact on the workplace.*
Technical digital skills:	the complexity of the technological application and how defined it is	• A few defined and simple technological applications • The defined technological applications are simple to use.	• The technological applications are less defined which implies a selection. • The technological applications are moderately complex to use.	• The technological applications are undefined which implies a selection. • The technological applications are complex to use.
Information processing digital skills:	the level of inference needed to complete the task, how well the information is defined, and how difficult it is to process	• A few defined and simple pieces of digital information to process	• A few defined and more complex pieces of digital information to process which requires a higher level of inference.	• Several undefined and complex pieces of digital information to process which requires a high level of inference.
Setting:	the environment in which the task is performed.	• Precise • Defined	• Less precise • Less defined	• Imprecise • Undefined

There are 24 questions for 8 tasks in the assessment. Scoring includes a complexity score (based on the Complexity Matrix, which is a combination of two complexity scales: open/closed and well defined/ill defined), a task description score (the numbers refer to the other complexity dimensions) and the skills required to successfully complete the task.

The psychometric properties of SRDC's DES assessment have not been studied, therefore its precision and external validity are unknown, though the construct validity is clear.

A sample question of SRDC's DES assessment

Assessment for Task 6: Finding and Selecting Information Online – Interactive Version

Question 1 (Task 5)

You need to get the names of two stores that sell office desks in Canada. Use "office desks" as your search words using a search engine such as Google. Find the websites for the two stores. Then exit the internet.

A. Web page first appears like the one below with search area to enter the search words "office desks." To get this part correct, the test taker must enter data in the correct place and hit the enter button.
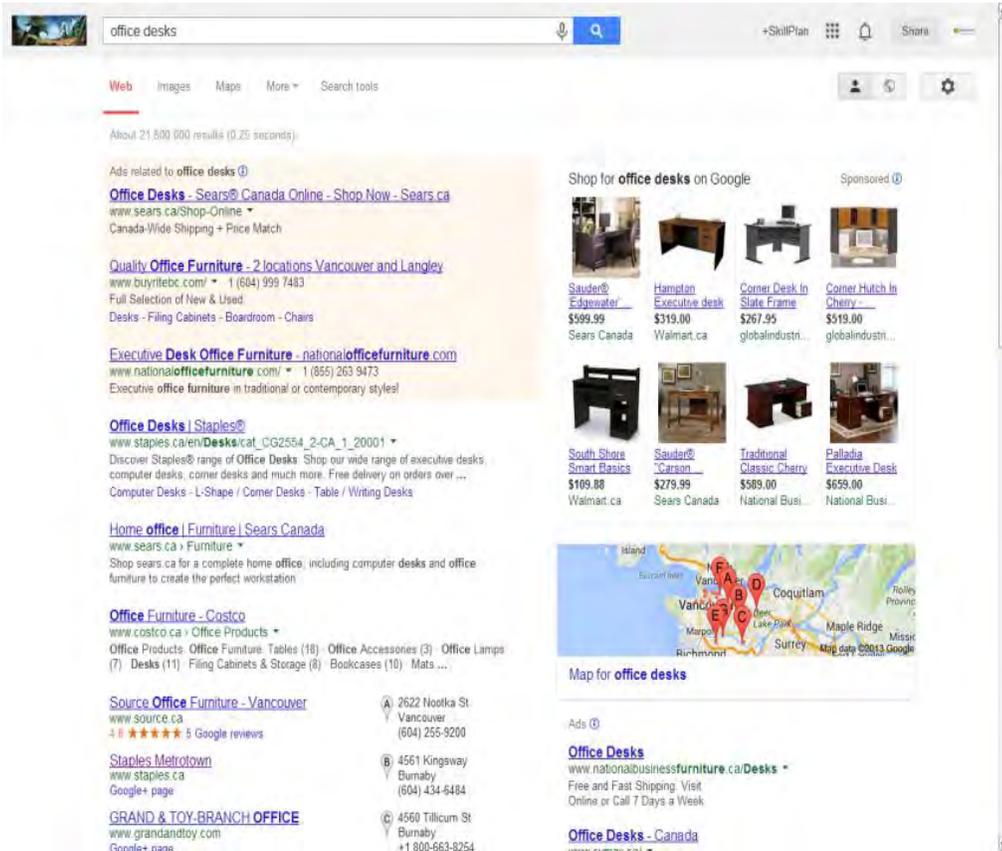
B. Click on one hyperlink and a web page for one store appears, like the following. Enter the store name in an answer box (not shown).
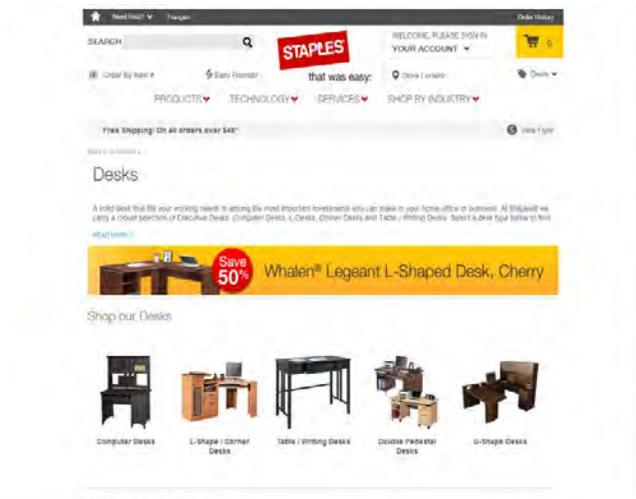
C. Click on a second hyperlink and the web site for another store appears like the one below. Enter the store name in the answer box next to the previous one (not shown).
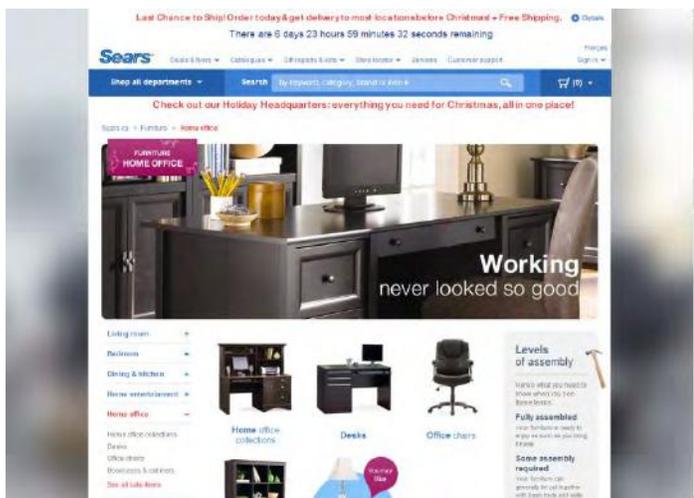
D. If the test taker has entered the names of two different furniture stores, then the question is considered correct.

Self-reported objective assessment – OLES Essential Skills Indicator

The Essential Skills Indicator[1] provides objective assessments of Numeracy, Document Use, and Reading for jobseekers and workers to administer themselves. The Indicator offers pre and post quizzes containing different sets of questions to help users measure their improvement after participating in a training program. The tests are divided into Levels 1, 2, and 3. Test-takers are recommended to begin with Level 1, but they can start with any level. The quizzes are not designed to be adaptive, that is, they do not stream test-takers into level-appropriate questions based on their performance in the previous section. Upon completing a test at a given level, test-takers can choose to move on to the next level above or below as they see fit. They also receive a result report at the end of every level they complete showing the questions they get right, the questions they get wrong, and the associated rationales behind the correct answers of all questions. Some sample questions of the Essential Skills Indicator are provided in Appendix B.

Other Essential Skills assessment unrelated to IALS

Communications and Math Employment Readiness Assessment (CAMERA)

The Communications and Math Employment Readiness Assessment (CAMERA) measures Numeracy, Document Use, Reading and Writing, as they align with Ontario's Literacy and Basic

---

Skills (LBS) framework. It is important to note that unlike other assessments described in the previous section, CAMERA is not aligned with the IALS Essential Skills levels (Preparatory Training Programs of Toronto, 2007). Therefore, results obtained from CAMERA assessments cannot be compared with results obtained from other Essential Skills assessments such as TOWES or ESG assessments.

The key advantage of the CAMERA is its user-friendliness. It is not intimidating to test-takers, which makes it particularly suitable for test-takers who have had negative experiences with schooling and testing, or who have been out of school for a long time and need time to adjust to formal testing. Furthermore, it is a well-integrated part of a comprehensive training system, which means the assessment is tightly aligned with training activities and highly relevant to both learners and trainers. Figure 13 provides some sample CAMERA tasks.

Figure 13   Sample tasks in CAMERA **–** Writing, Document Use

Source: The CAMERA Booklet. http://ptpcompass.ca/wordpress/wp-content/uploads/2016/06/CAMERA-E-Book.pdf

Self-reported assessments of the nine Essential Skills

Self-reported assessments of the nine Essential Skills are designed to be administered by the learners themselves to help them explore and gain a better understanding of their own skills sets. These tools are also meant to be integrated into users' learning paths, as they usually provide users with not only a snapshot of the current states of their skills and abilities but also suggestions on how to enhance their strengths and address areas of weakness. At the end of these assessments, users typically receive a template or a result report that encourage them to reflect on their current skills levels, with tips on how to better utilize their Essential Skills in their jobs.

By definition, these self-reported assessments are not customized to align with any single industry context. Instead, they are designed to be broadly applicable to a range of work settings. The intended users of these assessments are both jobseekers and employed workers in multiple occupations. In contrast to the objective assessments described in the previous section, the assessments outlined below are subjective in nature, that is they ask respondents to reflect on their own skills and report their own perceived ability/confidence to perform tasks that require these skills.

IALSS Self-assessed Indicators

The International Adult Literacy Skills Survey (IALSS), a later version of the IALS, has two sets of self-assessed questions measuring Reading, Document Use, and Numeracy skills. These questions ask about 1) one's _confidence_ in the application of these skills, as well as 2) the _frequency_ of their usage. SRDC has used these indicators to measure skill changes after skill development interventions (i.e., in the UPSKILL and Foundations projects), which fundamentally aim to increase not only the knowledge components underlying these skills but also their performance and the application in multiple contexts. Improved confidence in these skills is a good indicator of skill acquisition, while increase in frequency of skill usage is a strong proxy of changes in the underlying competency.

For instance, regarding the _confidence_ indicators, respondents are asked to rate on a 5-point scale the extent to which they agree or disagree with statements such as the following: "I have the reading skills in English I need to do my main job well" (Reading). With respect to the _frequency_ indicators, respondents are asked to rate on a 6-point scale how often they do activities such as the following as apart of their main job: "Calculate prices, costs, or budgets" (Numeracy).

Besides questions that refer to the workplace context, the assessment also asks about confidence and frequency of skill usage outside of work. To give some examples, for confidence indicators, respondents rate the extent to which they agree or disagree with statements such as "I feel anxious when figuring out such amounts as discounts, sales tax or tips" (Numeracy), or "Reading is one of my favourite activities" (Reading). For frequency of usage, the question specifies that respondents should answer how often they do activities such as the following outside of work: "Read or use information from newspapers or magazines" (Reading and/or Document Use).

Assessing the frequency of usage and confidence in application of Essential Skills outside of the workplace provides another piece of evidence of skill acquisition and competency. The underlying

assumption of skills training interventions is that the skills gained would be transferable to different contexts beyond the workplace. In cases where the period between baseline and follow-up assessments is not long enough to observe improvements in confidence and frequency of usage in the workplace, improvements outside of the workplace contexts could be valid proxies that illustrate skills gains. An example of a full set of questions from these kinds of assessments, as used in SRDC's UPSKILL project, is presented in Appendix C.

OLES self-assessments for jobseekers and workers

OLES self-assessments for jobseekers and workers[2] ask learners to reflect on their work and life experiences and provide a subjective evaluation of their abilities to perform a range of tasks. These tasks are intended to reflect common job responsibilities that workers tend to have in any workplace, regardless of industries or occupations. Statements used in these assessments closely reflect fundamental behaviours associated with each of the nine Essential Skills. For example, to assess their Numeracy skills, test-takers are asked to indicate on a 3-point scale of "Yes," "Somewhat" and "No" if they can "perform simple calculations such as addition or subtraction," or "record time using digital and standard clocks, watchers, or timers." The same format of assessment is used for the non-literacy skills such as Working with Others, Thinking and Continuous Learning. To illustrate, statements in the Working with Others assessment include "I can schedule and coordinate my work with the work of others," and "I can take initiative by doing what needs to be done before being asked."

At the end of each skill assessment, test-takers have an opportunity to complete a Personal Development section, which helps them document and reflect on their strengths (based on the "Yes" column) and weaknesses (based on the "Somewhat" and "No" column). The instructions users are given at the beginning of the assessments indicate that if they check the "Somewhat" and/or "No" columns for more than five indicators within a skill, they may want to consider upgrading that skill. This means users can use these tools to identify, in broad terms, the skill gaps they need to address, without any further details on their specific skill levels or contexts of their skills needs. Full assessments of all nine Essential Skills are included in Appendix D.

These assessments are designed to be applicable to multiple purposes. Users who want to enter the job market can use these assessments to explore the kind of basic skill requirements necessary to succeed in the workplace. Employees who wish to upgrade their Essential Skills can use these tools as a starting point to explore where they can improve. Overall, OLES self-assessments are simple tools for jobseekers and workers to gain a better understanding of their Essential Skills, but they are not designed with the kind of complexity and customizability that would allow for tighter alignment with either learners' training needs or employers' business needs.

---

[2]   See the Self-assessments section on https://www.canada.ca/en/employment-social-development/programs/essential-skills/tools.html

Ontario Skills Passport

The workplace-based assessments in the Ontario Skills Passport (OSP)[3] are designed to help workers and jobseekers learn more about Essential Skills and work habits, thus gaining a better understanding of the fundamental building blocks of a successful career. Skills in the OSP are categorized in a slightly different way than in the Essential Skill framework, but conceptually 15 of the skills and work habits in the OSP align well with seven of the Essential Skills. As shown in Table 8, Reading, Writing, Document Use, and Communication are defined in the same way in both the OSP and ESDC's Essential Skills framework. For Numeracy and Thinking, the OSP explicitly differentiates various aspects, dividing Numeracy into Money Math, Scheduling or Budgeting and Accounting, Measurement and Calculation, Data Analysis, and Numerical Estimation; and Thinking into Job Task Planning and Organizing, Decision-making, Problem Solving, Finding Information, and Critical Thinking. The OSP also defines Teamwork as one of the work habits that everyone in the workplace needs to have instead of as a skill with definitive levels of complexity on which people can progress.

Table 8 **Aligning OSP's Essential Skills with ESDC's Essential Skills**

ESDC's Essential Skills	OSP's Essential Skills and Work Habits	OSP's Skills Definition
Reading	Reading	The comprehension of text consisting of sentences and paragraphs
Writing	Writing	The preparation of written materials for a variety of purposes
Document Use	Document Use	The use of labels, lists, signs, graphs, charts, tables, forms, and other similar materials
Numeracy	Numeracy[4]	
	Money Math	The use of mathematical skills in making financial transactions, such as handling cash, preparing bills, and making payments
	Scheduling or budgeting and accounting	Planning for the best use of time and money, as well as monitoring the use of time and money
	Measurement and Calculation	The measurement and calculation of quantities, areas, volumes, and/or distances
	Data analysis	The collection and analysis of data in numerical form
	Numerical Estimation	The production of estimates in numerical terms

---

[3] See "Tasks I do or have done at work" and "Workplace tasks I think I can do" sections of
http://www.skills.edu.gov.on.ca/OSP2Web/EDU/SelfAssessmentTestHome.xhtml

[4] The OSP lists the 5 numeracy-related skills and the 5 thinking skills separately, but the behavioural constructs behind these skills are similar to that of **ESDC's Numeracy and Thinking skills.**

ESDC's Essential Skills	OSP's Essential Skills and Work Habits	OSP's Skills Definition
Oral Communication	Oral Communication	The use of speech for a variety of purposes
Thinking	Thinking[5]	
	Job task planning and organizing	**The planning and organization of one's own work**
	Decision-making	The making of any type of decision, using appropriate information
	Problem solving	The identification and solving of problems
	Finding information	The use of a variety of sources, including written text, people, computerized databases, and information systems
	Critical thinking	Making judgments by using criteria to evaluate ideas and information and the related consequences
Working with Others	Teamwork	Working willingly with others
		Showing respect for the ideas and opinions of others
		Taking responsibility for his or her share of the work
		Contributing

Two kinds of self-assessment for a workplace context are provided in the OSP. Generic assessments of workplace tasks ask people if they think they can carry out the tasks that are generally expected of employees in most jobs. We will focus on these generic assessment in this section, while revisiting the OSP in a later section that describes assessments that are contextualized to specific occupational skills requirements.

With the exception of Teamwork, the generic self-assessments of skills in the OSP show respondents a task and ask them to reflect: "Do I think I can do this?" An example of a task under Document Use is "Interpret data from a variety of graphs. For example, examine line graphs of hours worked and cost incurred. Interpret the graphs to identify patterns such as increased labour costs." The level of complexity of each task is given at the end of the description of the task. To illustrate, the Document Use task described above is a Level 3 task, according to the OSP. For each skill level, two to five tasks are included in the assessments.

Teamwork, which aligns with Working with Others, is assessed through a scale with multiple indicators. Respondents are asked to rate on a 4-point scale how well statements such as the following describe what they do at work: "I work co-operatively with my supervisor and co-workers to get the job done," or "I listen respectfully to the ideas of others, even if I don't agree with them." Unlike the other skills, no level of complexity is associated with Teamwork. Full examples of all the assessments are included in Appendix E.

At the end of the assessments, respondents receive a results report indicating their skills levels, based on how many tasks they think they can do within each level. The results report is a simple summary to help respondents quickly see a record of their assessments. It is up to the respondents to interpret the results, make sense of the skills gaps, devise a training plan to upgrade their skills, and re-assess their skills after training to record any improvements and/or evaluate the effectiveness of their training plan.

Assessments of industry-contextualized Essential Skills

Since the early 2000s, there has been a growing consensus in the field of adult literacy and essential skills training that approaches emphasizing "learning in context" and "learning to do" are more effective than the general academic approaches. Therefore, there has been a shift towards program models that deliver LES training in the context of industry-specific skills. A growing body of assessments has emerged as a result, measuring participants' skills improvements after training and documenting data for the evaluation of these programs. Compared to the generic assessments, a key difference in these assessments is that Essential Skills are contextualized as the underlying capabilities that workers utilize and apply in order to successfully carry out the day-to-day job tasks in their occupations. Typically, these industry-contextualized assessments are designed with a certain level of involvement from employers and industry stakeholders, whose input and feedback are incorporated in these assessments to reflect the skills needs of their workforce.

Assessments that are contextualized to the work contexts of specific industries can vary in terms of how tightly they are aligned with occupational skills needs. At one end of the spectrum, there are assessments that are built based on the general descriptions of job tasks in the Essential Skills profiles created by ESDC for more than 350 occupations. These assessments can be used by workers in these occupations at any company, regardless of size, growth trajectory, location, and other organizational characteristics. At another end of the spectrum, there are assessments that emerge from elaborate and involved process of Organizational Needs Analysis (ONAs), in which owners, managers and workers from multiple levels of seniority in the organization provide details on the business needs, performance gaps, skills gaps, and training needs of their workplace. The kinds of assessment that result from this process are typically more tightly aligned with the specific needs of the organization. Some of these assessments may even directly incorporate authentic workplace materials as part of the questions.

There are also industry-informed assessments that fall in between the two extremes of the spectrum, with a mix of specific and generic measures of Essential Skills embedded within measures of job tasks. One example of this type of assessments, which will be discussed in further details later, is a set of assessments developed for different industries from a common test bank. The test developers first come up with a bank of 45,000 questions measuring Essential Skills in generic terms. They then collaborate with different industries to understand the assessment needs of the industry users. Once sufficient industry input is collected, the test developers go into their test bank to select, modify, and customize the questions to introduce varying degrees of contextualization in the questions to meet the occupational requirements of the industries. Informed by industry, the resulting assessments have varying degrees of contextualization, uncovering performance gaps, skills needs, and training needs with different levels of details and specificity.

The rest of this section describes existing industry-contextualized assessments of Essential Skills that SRDC has reviewed through the environmental scan and through consultations with experts in the field. Two kinds of assessments will be discussed. First, instruments that assess respondents in an objective way, either through questions with definitively right and wrong answers, or through

expert observations using industry-standardized criteria, will be described. Next, we describe assessments that are more subjective in nature, with either self-assessed measures or employer-reported indicators.

Objective/expert assessments of industry-contextualized Essential Skills

Assessments sorted into this category are designed to objectively judge the skills levels of test-takers. Included in this category is the type of self-completed assessments in which respondents are asked questions with definitively right and wrong answers. Assessments in this section can also be based on objective opinions or informed impressions of expert examiners observing and judging individuals in their natural work environment. A key characteristic of these assessments is that they use well-defined, pre-determined answer keys or performance rubrics to evaluate the skills and performance of test-takers.

Assessments based on Bow Valley College's Test of Workplace Essential Skills (TOWES)

Besides the General Series of tests on Reading, Document Use, and Numeracy (described in an earlier section), TOWES developers have also worked with a variety of sectors and industries to create a set of proprietary assessments tailored to different industry contexts. To better understand these assessments, in conjunction with our own document review and environmental scan, SRDC has also conducted an in-depth interview with a representative from Bow Valley College (BVC) who has been closely involved in designing and testing these assessments (see Appendix I for a copy of the interview protocol). The interview sought to uncover how Bow Valley College developed these assessments, with a particular focus on the development of soft essential skills assessments such as oral communication, thinking skills, and working with others. We also discussed whether these measures have been field-tested and validated, and the contexts in which they could be used (e.g., a classroom setting, a workplace setting, before or after training, etc.).

Bow Valley College has worked closely with industry stakeholders and employers to develop Sector Series tests assessing the Reading, Document Use, and Numeracy skills needed for safe and productive employment within specific industries. These assessments incorporate authentic workplace materials to measure the extent to which test-takers have the competencies to carry out the daily job tasks required of typical workers in these industries. For example, the assessments for the manufacturing sector contain questions in the contexts of assembly drawings, schematics, as well as safety manuals and regulations. The Sector Series are also customized to the skill level requirements of these industries. For example, the assessment for the health sector targets skill levels 2 and 3, while the assessment for transportation professional focuses on skill levels 3 and 4.

To date, the following sector-specific assessments have been developed:

▪ TOWES Manufacturing (MFG): The MFG assessment measures skills at levels 2 and 3 and is best suited for machine operators, assemblers, or workers and factories and machine shops. The MFG covers Reading, Document Use, and Numeracy, and is designed to ask about assembly drawings, schematics, as well as safety manuals and regulations.

- TOWES Office and Administration (ADM): The ADM assessment measures skills at levels 2 and 3 and is best suited for clerks, receptionists, data entry personnel, and workers in business, non-profit organizations and public institutions. The ADM covers Reading, Document Use, and Numeracy, and is designed with the incorporation of manuals for office equipment, memos and office policies to contextualize its questions.

- TOWES Entrance into Apprenticeship (APR): The APR assessment measures skills at levels 2 and 3 and is best suited for pre-apprentices and apprentices in industrial trades. The assessment covers Reading, Document Use, and Numeracy, and includes measurement and calculation, reading and using work permits, schematics and codes.

- TOWES Aerospace (AER): The AER assessment measures skills at levels 2 and 3 and is best suited for workers in aviation services, maintenance and assembly. The assessment covers Reading, Document Use, and Numeracy, and includes job procedures, Material Safety Data Sheets (MSDS) and tooling references.

- TOWES Health Care (HCR): The HCR assessment measures skills at levels 2 and 3 and is best suited for Licensed Practical Nurses; Care Attendants; and Medical, Dental and Laboratory Technicians. The assessment covers Reading, Document Use, and Numeracy, and includes questions about blood pressure classifications, test requisition forms, and specimen collection manuals.

- TOWES Hospitality and Tourism Locator (HTC): The HTC assessment measures skills at level 2 and is best suited for those in the hospitality industry. It covers only two domains – Document Use and Numeracy – and includes questions on calculating services charges, reading about proper food handling procedures, and reviewing safety labels.

- TOWES Transportation Entrance (TRE): The TRE assessment measures skills at level 2 and is best suited for those entering into the professional driving training and industry. The assessment covers Reading, Document Use, and Numeracy, and includes questions that require respondents to determine vehicle loads and complete driver claim forms.

- TOWES Transportation Professional (TRP): The TRP assessment measures skills at levels 3 and 4 and is best suited for those completing professional driving training. The assessment covers Reading, Document Use, and Numeracy, and includes questions about reviewing engine reports and Overage, Shortage and Damaged (OS&D) reports.

In addition to these industry-specific assessments of Reading, Document Use, and Numeracy, the representative we interviewed confirmed that Bow Valley College has also developed a number of assessments for soft essential skills. The degree of contextualization to align these assessments with the skills needs of specific industries is substantial. According to the expert interviewed, Bow Valley College works closely with industry stakeholders to first identify the competencies – or observable behaviours and traits – that workers within these industries need to demonstrate on the job. For example, for Oral Communication, a competency could consist of a person's ability to clearly communicate a five-step process and use appropriate vocabulary. For each competency, a task or a scenario is then developed to assess a learner's skill level. These task assessments allow learners to demonstrate their skills and the results are then compared to the ES complexity

framework to give learners a rating and determine their competency level (e.g., a level 2 or 3 on Oral Communication). Since the objective of the exercise is to embed the assessment within tasks that competent workers need to execute in their jobs, these assessments are industry-contextualized.

Bow Valley College normally assesses learners through paper-based assessment instruments, except for skills related to oral communication and working with others. From Bow Valley's perspective, an individual's performance is tied to their knowledge, skills, and abilities. A paper-based assessment of oral communication and working with others is measuring their *knowledge* of these skills (e.g., what are the best practices of working cooperatively in a diverse team?), not their *skills* or *abilities* in a given context (e.g., how does an individual react when conflict between colleagues arises in the workplace?). To be able to assess observable behaviours in a given situation, Bow Valley College's assessments of individuals' oral communication skills and their ability to work with others are usually done by the LES instructor during classroom activities or classroom observations.

Bow Valley College's work in developing assessments for thinking skills has not been extensive, primarily due to the challenge of measuring these skills. In the limited work it has done, BVC has focused on assessing job-related thinking skills, such as problem solving, job tasks planning and organizing, and other similar subsets of skills, rather than general thinking skills. BVC applies a similar approach for developing assessments of thinking skills, first by identifying competencies, then creating tasks, and evaluating the ability of the learner to complete the task using a scoring rubric. Thinking skills are assessed through either a group activity or a paper assessment.

Continuous learning is typically not assessed by Bow Valley College for two main reasons. The first is due to the difficulty in creating an assessment that would be valid and would provide useful information. Indeed, according to BVC, continuous learning should be understood as "keeping pace with the changing demands of your job and the requirement to refresh your skills and knowledge." From that perspective, continuous learning falls within the literacy skills. However, a person's capacity to learn is highly related to their literacy skills. Hence, assessing a person's continuous learning skills requires looking at literacy.

The second reason is the lack of demand on the part of employers to assess continuous learning skills. Most employers tend to be prescriptive about the kinds of training that their employees should take, which means that the responsibility for making decisions about future training are removed from the employee and fall instead on the supervisor or the human resources department.

The fact that Bow Valley College has developed these assessments with substantial input from industries means that there is a high degree of contextualization and customization of these assessments to industry-specific skills needs. These assessment tools can be used by jobseekers participating in essential skills training programs with a well-defined employment pathway and/or target, with the goal to upgrade their skills to eventually obtain gainful employment in specific occupations within specific industries. They are also suitable for workplace training contexts for workers aiming to upgrade their skills to address performance gaps.

As confirmed by the key informant interviewed, these assessments are a resource that can be used at various points before, during, and after training programs. Trainers and program evaluators can use them to assess the baseline skills of participants before they go through the training program. They can be used at certain milestones during the program to document progress. They can also be used after the program to measure skills gains, which also serve as a tool to collect the necessary data to evaluate the effectiveness of the training programs.

Assessments based on Essential Skills Group's Essential Skills Assessments and Training

Besides the widely-used objective, non-contextualized tests of Reading, Document Use, and Numeracy described in an earlier section, the Essential Skills Group (ESG) has also developed a test bank with over 45,000 questions that can be contextualized to align with the work environments of various sectors, industries and occupations. In addition to collecting information from the document review, SRDC has also interviewed a representative from ESG to gain a better understanding of the process through which these assessments have been developed. According to our review and interview, the skills assessed by these tools go beyond the Essential Skills framework and can be grouped into the following categories:

- **Core literacy skills that are well-aligned with the Essential Skills framework**: Document Use, Numeracy, Reading, Writing, and Computer Use.

- **"Soft" skills that are slightly different in conceptualization to those in the Essential Skills framework**: Communication, Critical Thinking, and Teamwork, which have slightly different conceptualization than Oral Communication, Thinking, and Working with Others, respectively, in the Essential Skills framework. ESG has not developed any assessment of skills related to Continuous Learning.

- **Industry knowledge**: these assessments test the understanding of Business and Management, Customer Service, Security, Standards and Regulations, and Workplace Safety, which go beyond the Essential Skills framework.

Table 9 provides further details on the mode of administration of each of these skills. For Document Use, Numeracy and Reading, the structure and format of administration of these sector-specific assessments are similar to their respective general assessments. These assessments are composed of multiple choice questions with objectively right and wrong answers, and are adaptive in nature. Short versions of these assessments can have as few as nine questions for Reading and Document Use, and between 15 and 18 questions for Numeracy. The long versions are composed of 15 questions for Reading and Document Use and between 25 and 30 for Numeracy. These assessments have been adopted by various industries, including the supply chain industry and the trades sector.

Table 9   Skills domains a**ssessed by the Essential Skills Group's** tools

Skills	Mode of Administration
Core Literacy Skills	
Document Use	Use information contained in form, graphs and tables to answer questions.
Numeracy	Answer questions involving data analysis, equations, measurement conversions and the addition, subtraction and multiplication of decimals, fractions and per cents. People can use calculators.
Reading	Use information presented in workplace documents such as memos and bulletins to answer questions.
Writing	Answer three types of writing questions that assess basic grammar and writing skills.
Digital skills/ Computer Use	Use background knowledge of computer and digital skills to answer questions.
Soft Skills	
Teamwork	Use background knowledge of teamwork principles to answer questions.
Communication	▪ Speaking skills: Read statements and answer questions to show proficiency of speech ▪ Listening skills: Listen to audio clips and use the information they heard to answer questions.
Critical Thinking	Use information presented in scenarios to answer questions.
Industry Knowledge	
Business and Management	Use background knowledge of business and management principles to answer questions.
Customer Service	Use background knowledge of customer service principles to answer questions.
Security	Use background knowledge of security practices to answer questions.
Standards and Regulations	Use background knowledge of standards and regulations to answer questions.
Workplace Safety	Use background knowledge of workplace safety practices to answer questions.

Sources: Connector – A Skill Matching Tool User Guide, and SRDC interview with Michael Herzog.

The degree to which these assessments are contextualized vary from question to question. We have seen examples of questions on sector-specific assessments that are identical to those on the generic assessments, if the contexts of the generic questions happen to align well with the contexts of the industry of interest. Figure 14 presents a Document Use question that shows up in both the generic assessment and the assessment specifically designed for the trades. In this example, the question asks respondents to read a graph about pipe size and associated collapsing pressure and maximum operating pressure. While test-takers with no prior knowledge of piping can answer this question, it presents precisely one of the common work tasks that pipefitters and plumbers have to do in their jobs. It is therefore suitable to be adopted verbatim from the generic assessment into the industry-specific assessment.

Figure 14   Document use sample question

Answer the following question using the information provided.
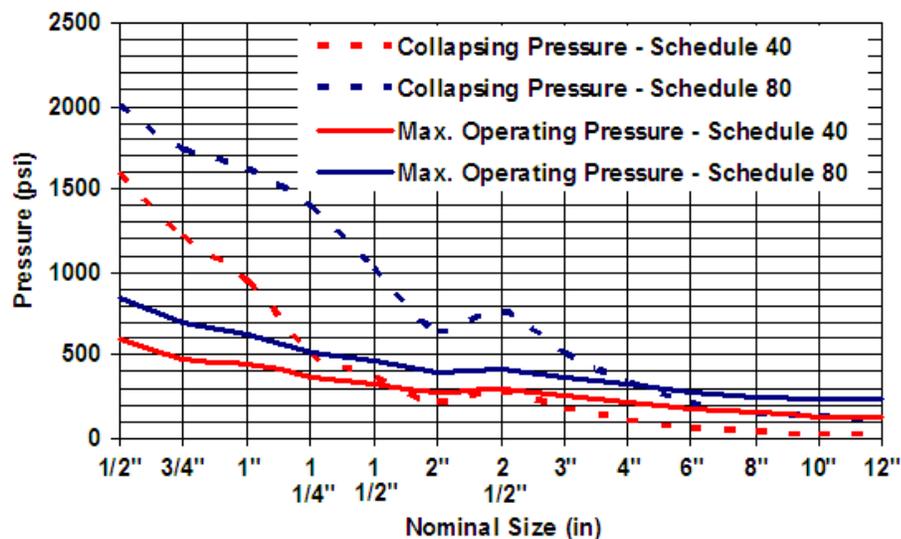
What size of Schedule 80 pipe has the same collapsing pressure and maximum operating pressure?

- o   2 inch pipe
- o   2 and 1/2 inch
- o   3 inch pipe
- o   4 inch pipe

On the other hand, there are questions that incorporate authentic workplace materials into its content, providing a substantially tighter alignment with the industry context. Figure 15 presents an example of such questions with high level of contextualization of trade occupations. To be able answer this Document Use question, workers in the trades must rely on their experiences using wiring diagrams to locate switches and power sources, a sector-specific skill requirement for the trades.

Figure 15   Document use sample question **–** with sector-specific document

Workers use wiring diagrams to locate switches and power sources. Use the schematic to answer the question that follows.
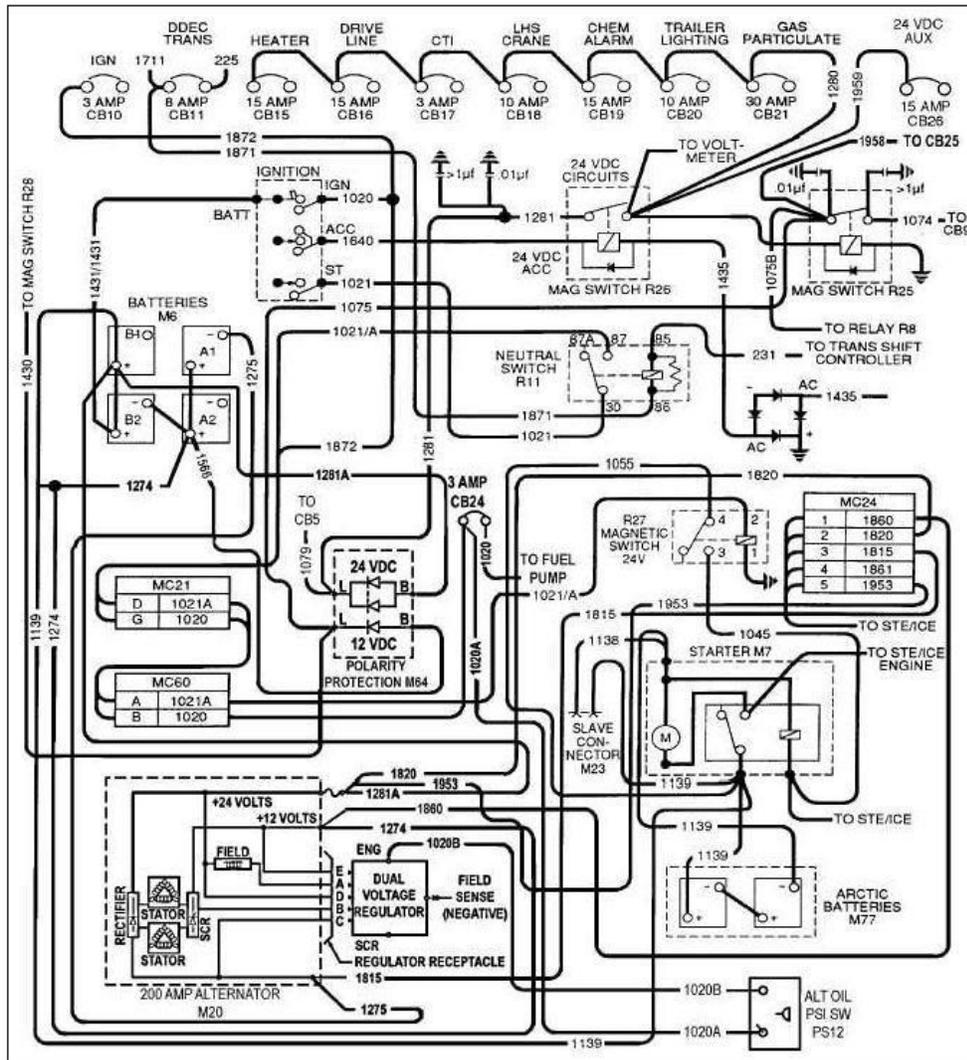
What component will no longer function if fuse CB24 is exposed to a 6 amp current?

o   Fuel pump
o   Oil pressure switch
o   Neutral safety
o   Starter

Source: ITA Essential Skills.

Other examples further illustrate the variability in the degrees of contextualization of questions in these assessments. During the interview, the expert gave an example of how they slightly adapted the context of a Document Use question to make it more suitable for specific workplace settings. The question presents a schedule of availability of six people. In the generic assessment, respondents are asked to think about these six people as family members trying to get together for a family reunion, determining the best time to host this event based on the given availability. To adapt this question for the workplace context, ESG test-makers modify the pre-amble of the question to present the schedule as availability of workers to work overtime. The question in this context then asks respondents that if one of the workers cannot make their shift, which of the remaining five workers are able to take over instead.

Another example of this varying degree of contextualization is in the Writing assessment. Questions in the Writing assessment ask respondents to spot and correct grammatical and spelling mistakes on pre-written documents. The content of the pre-written documents is tailored to reflect the working environments of different industries. Figure 16 provides a sample question in which the content is written specifically for drivers in the transportation industry.

Figure 16   Sample question – Writing (e-mails, memos, and letters)

In the e-mail below, select statements are underlined and identified by numbers in brackets, such as (1). These statements may have grammar and spelling errors. Use the e-mail and the underlined and numbered statement to answer the question that follows.

From: Jane Doe
Sent: Wednesday, Oct. 16, 2014 8:15 a.m.
To: All Transportation Drivers
Subject: Addition to Nov. 9 Meeting Agenda
Attachment: Safety Alert 13-001

Due to the recent increase in slip/fall incidents with drivers, Health & Safety will be adding an agenda item to the upcoming general meeting on <u>Nov. 9, 2014, at 10:00 am in the Great Hall. (1)</u>

<u>In preparation for this meeting please review the following two documents: (2)</u>

   1) Safety Alert 13-001 attached to this email

   2) <u>Mounting and Dismounting of Equipment found in Chapter 5 of the Driver Policy and Procedure Handbook (3)</u>

Your Health & Safety Committee <u>looks forward to working with you to address this important safety issue</u>.(4)

Jane Doe
Health & Safety Manager

What is the correct punctuation for statement (2)?

- o   In preparation for this meeting please review the following two documents,
- o   In preparation, for this meeting please review the following two documents:
- o   In preparation for this meeting, please review the following two documents:
- o   In preparation for this meeting please review the following two documents:

Source: Stocking Up on Skills.

Questions assessing Digital Skills or Computer Use tend to be less contextualized, as they ask respondents about the functionalities of computer, software, and other digital equipment commonly used by office workers across multiple industries. Figure 17 presents a sample question asking generally about word processing function.

Figure 17   Sample question **–** Computer use

Word processing software has a range of functions that allow users to compose and manage written material.

Which of the following is not a word processing function?

- o   Creating batch mailings
- o   Spell checking and auto-correcting text
- o   Creating tables of figures
- o   They are all word processing functions

Source: Stocking Up on Skills.

ESG's assessments of the Essential Skills framework's soft skills include measures of Teamwork (a subset of Working with others), Communication skills (specifically listening and speaking skills), and Thinking skills (including measures of decision-making, critical thinking, and problem solving). The interviews provided further contexts around the types of assessments used by ESG as well as their targeted populations and program purposes.

ESG developed an objective measure of Teamwork that assesses respondents' understanding of the importance of working with others and their knowledge of being an effective and productive team member. However, the representative from ESG told us that from the perspectives of employers and industry stakeholders, it is impractical to measure Teamwork skills objectively. This is because objective Teamwork assessments focus on respondents' knowledge of the principles of working in a team and not their actual abilities to coordinate work with others. The ESG expert acknowledged that the scenarios used in their objective Teamwork assessments are somewhat misaligned with the conceptualization of Working with Others in the Essential Skills framework, as they are not measuring actual behaviours, something that would be impossible to do using their online tools. For instance, this tool can inform an employer about whether their employee *understands* how to work with a team, but it does not assess whether the worker *can actually* work well in a team. Figure 18 provides an example of ESG's Teamwork indicator. Based on our experience with other instruments, the most useful way to assess Working with Others is to identify tasks linked with

underlying aspects of this skill in the context of Organizational Needs Analyses and develop self-reported measures about these tasks. We will discuss this approach further in subsequent sections.

Figure 18   Sample question **–** Teamwork

Answer the following question about teamwork.

On a high priority project, how often should a team meet?

- ○ As often as the team leader thinks is necessary in order to achieve the objective
- ○ As often as the project manager thinks is necessary in order to achieve the objective
- ○ It depends on the urgency of the decision
- ○ As often as the team members think is necessary in order to achieve the objective

Source: Connector.

ESG uses objective multiple-choice assessments and expert-assessed observations to assess two subsets of a learner's Communication skills: their listening and speaking skills. For Listening skills, respondents listen to a short audio clip with dialogue presenting details of a given scenario. Respondents must then answer questions to assess their comprehension of what they have just heard. As participants answer questions correctly or incorrectly, follow-up questions are either easier or more difficult. Figure 19 shows the layout and format of such Listening questions.

Figure 19   Sample question **–** Communication (Listening)

Listen to the audio clip about an accident and answer the following question.
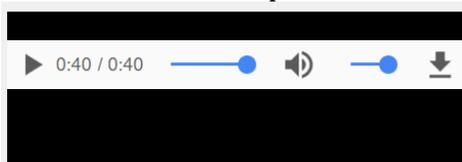
How long was the cut on the worker's leg?

- o   1 centimetre
- o   2 centimetres
- o   3 centimetres
- o   4 centimetres

Source: Connector.

For the Speaking component of the Communication assessment, respondents are first asked to read aloud statements of increasing complexity. Their answers are recorded and then sent to assessors

who use a rubric to measure the proficiency of the respondents' speech. The second part of this assessment provides respondents with a topic to assess their free-form speech. For example, a prompt for this part could be: "What is your favourite activity and why do you like doing that activity?" The prompts can be adapted so that respondents speak about their occupation or industry. Again, expert test-markers listen to recordings of the respondents' answers, looking for correct pronunciation, precise word choice, and appropriate pace and rhythm of speech. Overall, the Communication assessment is similar in content and format to language proficiency tests for foreign language learners.

ESG has developed roughly 54 multiple-choice questions that are currently being field tested to assess a learner's Thinking skills, which includes decision-making, critical thinking, and problem solving skills. For each of these questions, respondents are provided a scenario where a problem or an action is required (e.g., "*The forklift is teetering or overloaded*") and are asked to select from a number of options their preferred course of action to respond to that scenario. The questions are contextual, meaning that each of the scenarios provided are requirements of the workplace. Indeed, ESG commented in our interview that, in many ways, what is being measured through their assessment of thinking skills is common sense, but "common sense is only common if you've experienced the situation." As a result, for any assessment of problem solving or decision-making to reflect the effectiveness of a worker, the scenarios must reflect the realities of their workplace. Figure 20 provides an example.

Figure 20   Connector sample question **–** Thinking

Read the scenario and answer the question that follows.

Scenario

You recently were promoted to supervise a team of 10 material handlers with whom you have worked with for two years. Your manager asks you lay off one of the material handlers to reduce wage costs by 10 percent.

Which of the following is the most important consideration when deciding what to do?

- The newest of your role
- What the alternatives are
- The relationship with your staff
- The relationship with your supervisor

Source: Connector.

Overall, ESG began the assessment development process by establishing a substantive question bank of 45,000 items, which enables them to subsequently go into the bank, select and modify questions to fit with the varying contextualization and customization needs of industry users. For example, the needs of the Industry Training Authority (ITA) overseeing British Columbia's training

and apprenticeship system of more than 50 skilled trades is different from the needs of one employer wishing to measure their employees' skill acquisitions after a company-wide training program. For organizations such as the ITA, the assessments are contextualized to the trade sector but are designed with sufficient level of *generality* to be relevant to a wide range of trade occupations. For training programs customized to a smaller group of closely-related occupations, ESG works with employers, trainers and industry stakeholders, collecting input to respond to *specific* performance gaps and align with *specific* skills needs of their workforce. ESG have also worked with industry users that require a mix of generality and specificity in their assessments, customizing their assessments to align with different training contexts of different industries.

In general, ESG has collaborated with various organizations to provide customized assessments according to the needs of workers and learners in specific workplaces, industries and provinces. The following list of tools illustrates the application of ESG assessments in a wide range of contexts:

Stocking Up on Skills: Conestoga College Institute of Technology and Advanced Learning works with supply chain employers, other colleges and the Canadian Supply Chain Sector Council to address the skills gaps being experienced in the supply chain. The Stocking Up on Skills web application is part of this initiative and includes assessments comparing people's skills to the requirements of four supply chain occupational groups: warehousing and distribution, procurement and asset management, transportation and logistics, and planning and management. Each occupational grouping has its own assessments customized to its occupational needs based on input from employers and industry stakeholders. The following seven Essential Skills are tested: Reading, Document Use, Numeracy, Thinking skills, Writing, Digital Technology, and Listening skills.

Connector: Also designed for the supply chain sector, the Connector provides employers, learning system providers, sector associations and individuals with an online tool that helps identify occupation-focused strengths and skill gaps, connecting test-takers to a wide variety of education and training resources. The assessment tools are useful for employers, employees as well as job seekers. Employers can use the assessments and learning activities to gauge their employees' skills and make decisions regarding workplace training investments. Employees can use the tools to assess and upgrade their skills in relation to current or desired occupations. Job seekers can also use the website to explore careers, assess their strengths and weaknesses, and determine where to go to upgrade their skills.

Essential Skills for Employment and Education (ESEE): Funded by the College Sector Committee for Adult Upgrading, the ESEE web application features assessments that compare the Reading, Document Use, Numeracy skills of test takers to general academic and vocational requirements. The tool also includes more than 140 embedded American Sign Language (ASL) videos to support deaf and hard-of-hearing learners.

Essential Skills for Ontario Tradespeople (ESOT): Funded by the College Sector Committee for Adult Upgrading, ESOT web application serves as a tool to help Ontarians develop the Essential Skills they need to succeed at a wide range of Red Seal trades. The tool features assessments that compare the Reading, Document Use, and Numeracy skills of test takers to the requirements of 53 Red Seal trades. Each trade has its own suite of customized assessments.

Essential Skills for the Health Sector (ESHS): Funded by the College Sector Committee for Adult Upgrading, the ESHS web application helps Ontarians develop the Essential Skills they need to achieve employment and occupational success in the healthcare sector. The tool features assessments that compare the Reading, Document Use and Numeracy skills of test-takers to the requirements of 11 healthcare-related occupations. Each occupation has its own suite of customized assessments.

ITA Essential Skills: The Industry Training Authority (ITA) is the provincial crown agency responsible for overseeing B.C.'s industry training and apprenticeship system. This ITA web application can be used by anyone for free; however, it is specially designed for women, immigrants and Aboriginal people who are involved with ITA Labour Supply Initiative programs funded through the Labour Market Agreement. The website provides access to built-in and customized Essential Skills profiles for more than 50 skilled trades, as well as assessments that help determine how ready people are for apprenticeship training.

Build Your Career with Essential Skills: This tool is developed in collaboration with Employment and Social Development Canada's Office of Literacy and Essential Skills. The web application provides updated Essential Skills profiles for 50 high-demand occupations as well as assessments that are specific to each of these occupations. These assessments are adaptive and allow test takers to self-select the skills they wish to assess. At the end of the assessments, test-takers receive customized result reports that highlight their skills and abilities. The occupational groups targeted by this tool include those in healthcare, accommodation and food services, telecommunications, construction, transportation, manufacturing, green initiatives and agriculture.

Alberta Workplace Essential Skills (AWES): The AWES web application is designed as a tool to help employers conduct needs assessments to understand the processes, skills and culture of their organizations. It also provides tools to support initiatives that integrate workplace Essential Skills training into existing organizational materials and operations, including Essential Skills assessments that evaluate the success of these training initiatives.

Essential Skills… Your Career Starts Here: This is a tool funded by the Aboriginal Skills Group (formerly known as Vancouver Aboriginal Skills and Employment Partnership), whose mission is to increase Aboriginal employment in strategic markets. The society collaborates with First Nation communities, Metis Provincial Council and, importantly, with industry and training institutions. The web-based application provides trainers and clients with access to customized Essential Skills profiles covering a wide range of occupations. It also provides assessments that help determine the Essential Skills readiness of clients. In addition, at the end of the assessments, test-takers receive test result reports that highlight their skills and abilities, as well as a learning plans that point trainers and learners to the resources that can be used to address skill gaps.

Essential Skills Readiness Assessment: This web application is funded by Workplace Education Manitoba (WEM), featuring updated Essential Skills profiles for 57 high-demand occupations as well as assessments specific to each occupation. It focuses on occupations in the following sectors: healthcare, food services, telecommunications, construction, transportation, manufacturing, and agriculture.

Though some of these assessment tools were designed to be tightly aligned with the needs of specific programs, sectors, and/or target populations, others were used to assess general progress for a diverse set of learners in a variety of training contexts. Box 1 illustrates some of the challenges associated with this kind of large-scale implementation, particularly when assessments intended to be interpreted at a group level are used as definitive indicators of individual competency by some practitioners and learners.

Box 1   Challenges associated with using a single type of assessment tool (ESEE) across multiple program contexts: the Learner Gains Research Project

Started in 2012 by the Ministry of Advanced Education and Skills Development (MAESD), the Literacy and Basic Skills (LBS) program is designed to help adults in Ontario develop and apply communication, numeracy, interpersonal and digital skills to achieve a wide range of goals. These goals can include successful transitions to secondary school, apprenticeship, post-secondary education, employment, as well as increased independence in general. The program aims to align with the training needs of various populations, streaming participants into sub-programs designed specifically for Anglophone college students, Anglophone secondary school students, Anglophone community-based participants, Francophone college students, Francophone community-based participants, Indigenous individuals and participants who are deaf or hard-of-hearing.

In 2015, the MAESD expressed an interest using a formal Essential Skills assessment tool to gain a better understanding of **participants' experiences and progress through the LBS program. As a result, the Learner Gains Research Project was** developed and the Essential Skills for Employment Education (ESEE) assessment tools were used, measuring LBS **participants' Reading, Document Use, and Numeracy skills. More than 2,800 learners across 45 training sites registered for** the assessment and, of this group, 2,782 completed one or more of the assessments for a total of 6,563 assessments (Clark, Hennessey, & Herzog, 2016). Tests taken include 5,537 entry or pre-tests and 1,026 exit or post-tests during the period of the pilot project, resulting in a large sample for analysis.

Key findings of the Learner Gains Research Project illustrated that there were measureable skill gains achieved by clients across all LBS program streams, ranging from 11-point to 21-point increases on average across all three skills (Clark et al., 2016). Indigenous clients demonstrated the most substantial gains, with an average of 37-point increase in each skill. Community-based participants in both the Anglophone and Francophone streams also showed substantial gains of about 12 to 13 points for each skill.

Despite these results, the ESEE assessments were not well-received in the field and the Ministry actually decided to discontinue implementing them as a tool to assess learner gains moving forward. The analysis reported focused on the overall group results, which were positive, without taking into consideration the negative experiences that participants had on an individual level. Multiple program staff participating in the Learner Gains Research Project found it challenging to motivate learners to complete the assessments, as well as to help them interpret their individual test results, as Dr. Christine Pinsent-Johnson documented in her blog posts.[5]

---

[5]   See ***ESEE come, not so ESEE go… What when wrong with the Learner Gains Research Project
and what can we learn from it,*** April 2018, for a summary of the analysis.
https://policyproblems.wordpress.com/2018/04/18/esee-come-not-so-esee-go-what-went-wrong-with-the-learner-gains-research-project-and-what-can-we-learn-from-it/

To be fair, the ESEE tools are designed to be used for low-stakes assessment. The results may be used with other information to inform decisions that affect individual outcomes (e.g., need for training, job performance). They may also be interpreted in aggregate form to inform decision-making about large groups of people. The tests are not designed with sufficient precision to be the only tool informing high-stakes decisions at the individual level, such as using the test results alone to determine a given **individual's skill mastery.**

Unfortunately, this low-stakes nature of the tests was not communicated well to learners, leading to the misunderstanding that **if a learner achieved a score lower than the normal range on the test, it must mean they have "failed" the test and did not** have the skills needed to be successful in most jobs. According to Dr. Pinsent-**Johnson's interviews with program staff,** learners were put in an untenable situation as they received discouraging messages indicating personal failure and inadequacy after spending substantial amount of time and efforts on the assessments. The majority of these learners have had negative experience with the education system throughout their years of schooling, many of whom were using the LBS program as a last resort to help them transition into the labour market. Such misleading and negative interpretation of individual scores on these Essential Skills assessments seriously affect their morale. Some learners refused the take the post-test after having unpleasant experiences with the pre-test. Some programs even lost learners – after taking the pre-test, participants simply left the program. Those who stayed have shared mostly negative responses with instructors and coordinators, including frustration, incredulity and tears. One instructor said some of the learners she worked with were **"decimated" after taking the test.**

The negative experiences of these learners could have been avoided if better messaging and implementation strategies had been in place. One lesson learned from this project is that practitioners on the ground should have been better equipped with the right messaging tools and resources to communicate the purpose and meaning of these assessments appropriately to learners. For example, it should have been emphasized that the test results are most meaningful when interpreted in aggregate to show progress of all participants as a group. On an individual level, the results are indicative and not definitive, with higher scores *suggesting* higher likelihood of successfully completing the training and moving closer to their education or employment goals.

More importantly, this example illustrates that it is best to incorporate assessments as part of a broader performance framework with multiple potential indicators of success. For example, when the Essential Skills assessments were administered in the Pay for Success project, participants also completed short self-assessments of other employability skills including career planning, career decision-making self-efficacy, job search clarity, and job search self-efficacy. When program coordinators and instructors helped participants interpret their individual test results, the Essential Skills results were interpreted in the context of these other employability skills, some of which were as predictive or more predictive of success in subsequent milestones along the employment pathway. Thus program staff were in a good position to use multiple indicators, without weighing one more heavily than the others, to provide tips and recommendations to ensure participants made the most out of the training and enhanced their chances of success. This helped participants understand the importance of Essential Skill assessments, without discouraging them if they received scores that were lower than expected.

SkillPlan's Self-directed Assessment – Measure Up

SkillPlan has developed a tool to help learners explore different Essential Skills requirements related to Numeracy, Document Use, and Reading for a variety of sectors, such as accommodation and tourism, construction trades, and trucking services. These are short assessments with objective questions to help learners visualize the types of questions they would encounter on objective

Essential Skills assessments, as well as the types of tasks they would need to resolve on the job. Figure 21 shows the options learners can choose to customize their assessment.

Figure 21   **Measure Up's st**arting screen
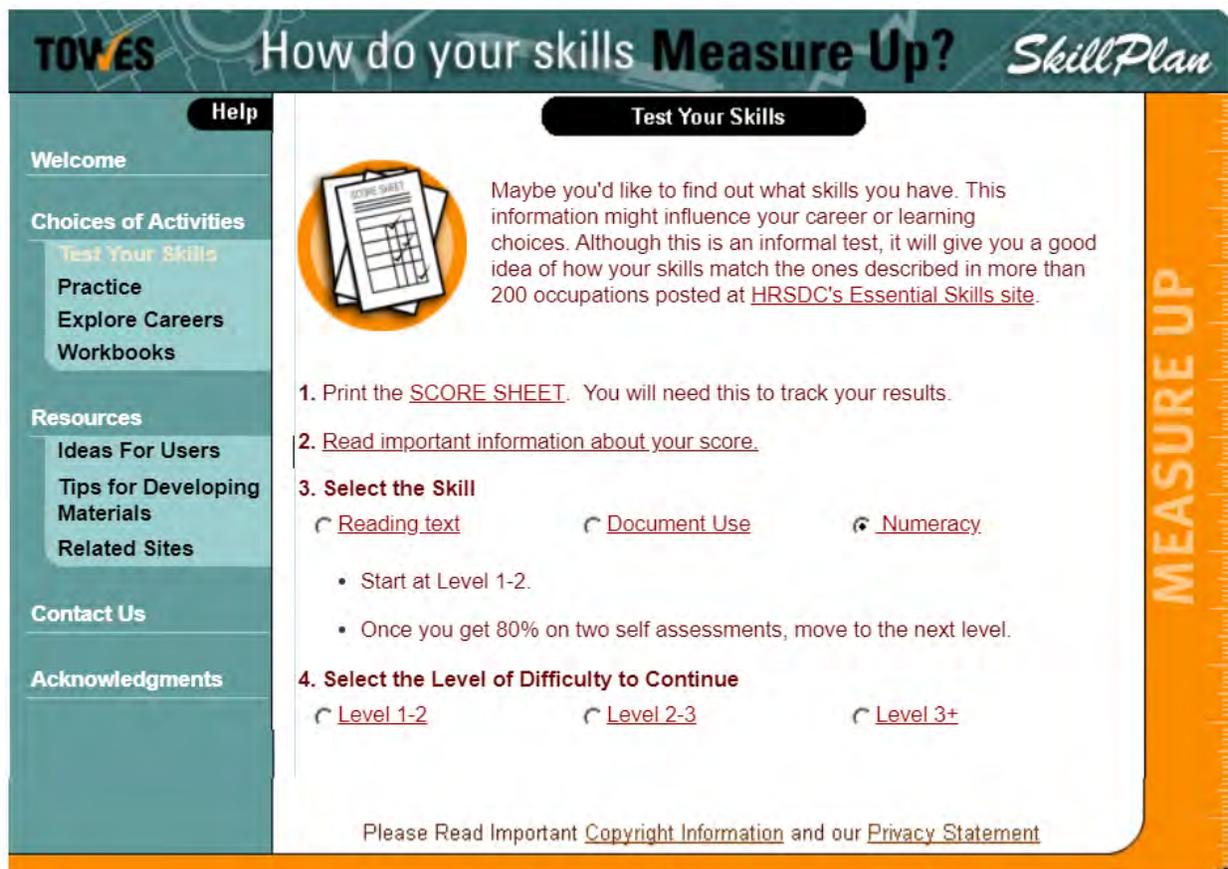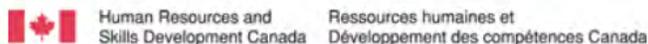
As shown in Figure 21, learners can start by downloading and printing the score sheet. After reading important information about their score, they can select the skill they want to test. The recommendation is to start at Level 1-2, and move to the next level once they get 8-% or more on at least two self-assessments. The levels of difficulty available are Level 1-2, Level 2-3, and Level 3+.

Figure 22 shows an example of a Numeracy task, at Level 2-3.

Figure 22   Measure Up – Numeracy – Level 2-3

On their own, these assessments can serve as a supporting tool to help learners and trainers check in and gauge learning progress on an informal basis. Beyond this purpose, these assessments provide a solid starting point to help project staff develop and refine contextualized Essential Skills measurements that are more appropriate for program evaluation.

SRDC's project – UPSKILL

UPSKILL is one of the pioneer studies that demonstrate the substantial, positive impacts that workplace literacy and essential skills training has on workers' skills, job performance, and a range of economic and social outcomes for workers and firms. The assessment tools used in this study include both standardized skills assessments of Document Use and Numeracy and observational industry performance assessments linked with National Occupational Standards. In addition, as mentioned in an earlier section, subjective, self-reported measures of Essential Skills usage in both the workplace and the everyday life contexts are also included. For the standardized assessments, the Test of Workplace Essential Skills (TOWES), which was described in a previous section, was

used. This section provides further details on the integration of Essential Skill assessments into pre-existing observational industry performance assessments.

Before these assessments are discussed, it is important to introduce the theoretical foundations behind the UPSKILL training model. Targeted to the tourism industry, this model evolve directly from the key business priorities of tourism companies. The UPSKILL project team conducted Organizational Needs Analyses (ONAs) and consulted with owners, managers, supervisors, and workers at participating businesses to unravel their business problems to the root causes. They then identify the learning components of these causes, and developing a set of learning requirements and a clear understanding of how the root causes will be solved through this learning. A performance framework emerged from this process, aligning the business needs of the employers with their employees' performance gaps and essential skills needs. Business needs are the business goals and objectives the organization must achieve to be successful (e.g., customer satisfaction, revenue increase, etc.). Performance needs are on-the-job accomplishments and behaviours that are required of employees in each of the target occupations in order to contribute to the achievement of these business goals. Performance needs identify what individuals must do more, better, or differently if the business goals are to be achieved. Performance needs are usually described in behavioural terms. Through these ONAs, causes of gaps in performance were also uncovered. Causes of gaps in performance may include: factors that are external to the organizations such as increased competition; factors that are internal to the organizations such as a lack of clarity about roles, a poor incentive system or inefficient work processes; and factors internal to individuals such as essential skills gaps. The training program of UPSKILL was designed to address these internal factors, aiming to enhance participants' essential skills, which translate to performance improvements that can have direct impact on enhancing the underlying business outcomes for the employers.

The measures of Essential Skills embedded within the workplace performance assessments are based on emerit©, a pre-existing industry certification performance assessment coordinated by the Canadian Tourism Human Resource Council (CTHRC). The original assessments are updated and modified to better align with the Organizational Needs Analyses results and to sufficiently cover Essential Skills concepts such as Oral Communication, Working with Others, and Thinking. The assessments are further contextualized to specific occupations within the sector, and target the following positions: custodial positions (housekeeping room attendants, maintenance), administrative and sales positions (front desk agents, reservation sales agents), serving occupation (food and beverage servers, banquet servers), and culinary occupations (line cooks).

The primary goal of these assessments is to measure employees' performance in key areas. Some of the performance areas have direct, one-to-one linkages with Essential Skills. For example, one of the key performance standards required of food and beverage servers is the ability to use effective verbal communication. To assess this, expert assessors observe employees in their working environment and check if they meet the behavioural standards required. These include making eye contact with speaker; showing interest in what is being said, for example, leaning forward slightly, avoiding interrupting, acknowledging with nod or smile as appropriate; and confirming understanding, e.g., repeating message in own words, ask questions to clarify. The evaluation

criteria used here aligns well with the conceptualization of Communication skills in the Essential Skills framework.

Some performance areas align with multiple closely-related Essential Skills. For example, front desk agents are assessed on their ability to respond to guest's concerns and complaints, which requires strong Communication and Thinking skills. For this assessment, the assessor interviews the employee and asks an actual time when the employee dealt with an angry guest who complained about his/her room or hotel staff. The assessor can also present a hypothetical scenario in which a guest complains to the front desk agent about dirty bed sheets after all housekeeping staff has finished their shifts for the day. In both cases, the assessor looks for evidence of behaviours that align well with both Communication and Thinking skills in the Essential Skills framework, such as the ability to acknowledge concern, apologize for inconvenience, thank guest for voicing concern (Communication); or the ability to select appropriate solution and request management assistance if unable to handle situation or satisfy guest (Thinking).

Finally, the assessments also include performance areas that require a cross-sectional mix of distinctive Essential Skills from the employees. One example of such performance areas is the ability to effectively promote the property, its services, and local tourism. Customers will regularly seek this information from any available staff, and making sure that all staff at the establishment are able to accurately provide such information is essential to the hotel's success in customer satisfaction. A positive exchange increases the likelihood of both further purchases and return stays at the hotel. Moreover, the information must be accurate and current to be useful, requiring staff to read and use documents to stay up-to-date on local events. Thus, this performance area requires a variety of Essential Skills from employees, including reading, document use, thinking, and communication to locate, understand, and convey the requested information. In the assessment of this performance area, the assessor asks the employee about a local event or a type of restaurant he/she would like to visit, taking note of the employee's ability to answer as well as the amount of enthusiasm conveyed. Evaluation criteria include the ability to 1) share knowledge and pride of local area; 2) provide information as requested; 3) direct guest to other information services for additional information; and 4) create favourable impressions that encourage guest to stay longer, use more information and services during stay, and visit again.

Overall, assessments that emerge from this type of training program models achieve a tight alignment between Essential Skills and job performance standards of employees as well as key business priorities of employers. The development of such Essential Skills measures require a deep level of understanding of the occupations, the organizations, the industry, as well as any external economic factors influencing both the employees' ability to meet performance standards and the employers' ability to achieve targeted business outcomes. This not only makes the assessments incredibly relevant applicable to the specific occupations under examination, but also takes the conceptualization of skills within the Essential Skills framework to a level of practicality beyond what is achieved by other types of assessments reviewed thus far.

Self-or-employer-reported assessments of job performance, based on underlying Essential Skills

Unlike objective or expert-evaluated assessments described in the previous section, these assessments rely on the subjective impressions of either the test-takers or their employers to gauge their Essential Skills. The instruments that collect test-takers' self-assessed evaluation reflect their self-efficacy in applying Essential Skills to carry out job tasks effectively. The employer-reported assessments ask supervisors and managers to provide their informed opinions on how well their employees or prospective employees perform on the job using their Essential Skills. A key characteristic of these assessments is that they rely on subjective judgment to measure the individual's willingness and ability to apply Essential Skills to complete job tasks.

OLES self-assessments for the trades

OLES provides informal assessments for apprentices and journeyworkers to gauge the levels of Essential Skills required for a career in the trades. These assessments ask learners to reflect on their work and provide a subjective evaluation of their abilities to perform trade-related tasks. Statements used in these assessments closely reflect the fundamental behaviours of each of the nine Essential Skills, but unlike the generic assessments presented in the previous section, these statements are contextualized in a trades work environment. For example, to assess Document Use, test-takers are asked to indicate on a 3-point scale of "Yes," "Somewhat" and "No" if they are confident in their ability to "understand labels, such as shipping or hazardous materials labels (e.g., WHMIS)." The same format of assessment is used for the non-literacy skills including Working with Others, Thinking, and Continuous Learning. The trades context is sometimes provided as example further illustrating the job task in question. For instant, statements under the Continuous Learning assessment include: "I am confident in my ability to learn from others (e.g., seek feedback from an experienced journeyperson)." At the end of the self-assessments of all skills, test-takers are presented with a Skills Summary section that provides them with a guideline to identify their essential skills strengths and areas for improvement. Full assessments all of nine Essential Skills are provided in Appendix F.

Ontario Skills Passport

Besides the general assessments described in an earlier section, the Ontario Skills Passport (OSP)[6] also provides self-assessed tools that are customized to the skills requirements of specific occupations. For these assessments, the OSP closely follows the Essential Skills profiles that ESDC has built for about 350 occupations in the National Occupational Classification (NOC) system. To assess each of the skills, the OSP asks test-takers to check off tasks they have done related to that skill. For example, for the Reading assessment, dental assistants are asked if they have "read reference books, e.g., read oral pathology reference books to learn about diseases of the gums and

---

[6] See "Tasks I do or have done at work" and "Workplace tasks I think I can do" sections of
http://www.skills.edu.gov.on.ca/OSP2Web/EDU/SelfAssessmentTestHome.xhtml

teeth," or "read manuals and guidelines to learn about methods and procedures for your work, e.g., read guidelines published by the Canadian Dental Association to learn about oral care for older adults." The format is similar for the assessments of the "soft" skills as well. To illustrate, tasks that dental assistants are asked to check off to assess their Communication skills include "exchang[ing] information with dentists, e.g., listen[ing] to the dentist's directions before and during examinations and discuss patient treatments and conditions," and "talk[ing] to patients and their relatives, e.g., explain[ing] treatments and oral hygiene practices to young patients and their parents." Appendix G provides examples of all nine assessments for dental assistants (NOC 3411).

SRDC's project – Innovative Cluster Training Model for New Brunswick Tourism Business

Building on the success of earlier work including UPSKILL, the *Innovative Cluster Training Model for New Brunswick Tourism Business* project was developed to provide small and medium-sized enterprises (SMEs) with a cost-effective training solution. Small businesses, particularly those with fewer than 20 employees, often face significant challenges in providing training to their staff. They usually lack the capacity and infrastructure to support on-site training delivery, or they are not in the position to provide significant release time for staff to attend off-site classes. To address this challenge, this project adapted existing best practices to construct a training model that takes into account operational and capacity constraints facing small businesses, aligning training delivery with their unique skills development needs. The overarching process used to develop both the training curricula and the assessment tools is similar to UPSKILL. Specifically, the project team conducted Organizational Needs Analyses (ONAs) to confirm and validate the alignment between business needs, performance gaps, and skills needs. However, since this training program is designed to address the unique challenges and needs of small businesses, these materials, including Essential Skills assessments, address one additional layer of complexity. They were specifically designed to respond to the need to pool learners and firms to achieve economies of scale and make learning more accessible for small businesses, while maintaining the ability to customize training and achieve tight alignment with unique learner and business needs.

Furthermore, staff in small firms are more likely to play the generalist role rather than specializing in single-faceted skill areas. They are more likely to "wear multiple hats" at a time and carry out a variety of tasks that requires multi-skilling on a daily basis. As a result, each of the Essential Skills measures embedded within job performance indicators emerged from this project often covers more than one Essential Skills. They are also generalizable enough to be applicable to a wide range to small tourism businesses, and at the same time specific enough to align with the occupational requirements of participating employees.

The assessment tools for this program are designed with a great deal of input from employers and industry stakeholders. Particularly, the project team works closely with the tourism industry in New Brunswick to come up with a set of job tasks that workers in tourism businesses complete on a daily basis. The assessment tools are then designed to measure Essential Skills in the context of performance on these specific job tasks. These assessments are self-reported and ask respondents to rate the extent to which they agree or disagree that they have all the skills needed to do their job well in each of the key performance areas. For example, under Customer Relations, some performance indicators align well with the behavioural requirements of both Communication and

Thinking Skills, such as "listening to and understanding customer needs and request," "speaking clearly and appropriately to address customer needs or requests," and "recognizing and responding to non-verbal expressions of customer dissatisfaction or discomfort." Table 10 provides the indicators and their alignment with Essential Skills. The full, formatted assessments of Essential Skills in the context of job performance used in this project is provided in Appendix H.

Table 10   Essential Skills indicators in the context of job performance assessment

I have all the skills needed to do my job well in the following areas:	Reading	Writing	Document use	Numeracy	Digital Literacy	Oral Communication	Thinking Skills	Working with Others	Continuous Learning
Understanding how my daily tasks contribute to meeting customer expectations.			✓	✓	✓	✓	✓	✓	
Understanding how the way I interact with customers contributes to business success.						✓	✓		
Having a positive attitude when dealing with customers (e.g., being willing to adapt and be flexible in a difficult situation).						✓	✓		
Taking initiative when dealing with customers (e.g., making decisions independently; doing things before being asked).						✓	✓		
Listening to and understanding customer needs and requests.						✓	✓		
Speaking clearly and appropriately to address customer needs or requests.						✓	✓		
Recognizing and responding to non-verbal expressions of customer dissatisfaction or discomfort.						✓	✓		
Asking clarifying questions to confirm information received from customers (e.g., credit card information, reservation details, etc.).						✓	✓		
Showing interest with non-verbal cues (e.g., eye contact, stopping what you're doing) when talking to customers.						✓			
Responding to email requests or enquiries.						✓	✓		
Listening to and understanding customer complaints.						✓	✓		
Asking clarifying questions to confirm customer concerns.						✓	✓		
Negotiating and conflict resolution (e.g., de-escalating conflict with customers).						✓	✓		
Exploring options to resolve complaints, and deciding on a solution.						✓	✓	✓	
Documenting complaints and resolutions according to standard procedures.			✓		✓				
Knowing how and where to find information to solve problems that come up on the job.			✓		✓		✓	✓	✓

I have all the skills needed to do my job well in the following areas:	Reading	Writing	Document use	Numeracy	Digital Literacy	Oral Communication	Thinking Skills	Working with Others	Continuous Learning
Collaborating with co-workers to solve problems.						✓	✓	✓	
Dealing with tool and equipment problems.			✓		✓		✓		✓
Going 'above and beyond' basic service requirements to create experiences that exceed customer expectations.						✓	✓		✓
Providing service that encourages first-time customers to return.						✓	✓		
Knowing and being able to talk about my business's services and facilities.						✓	✓		✓
Knowing and being able to talk about the community (e.g., restaurants, cultural activities, transportation options) in which my business is located.					✓	✓	✓		✓
Planning and organizing job tasks to complete high-priority tasks on time.							✓		
Reducing inefficiency (e.g., spending too much time on low-priority tasks).							✓		
Performing key job tasks accurately and being able to spot errors.							✓		
Coping with distractions while performing key job tasks.							✓		
Completing job-related documents accurately.			✓						
Understanding the roles and responsibilities of my co-workers.								✓	
Providing feedback to, and accepting feedback from, my co-workers.								✓	✓
Offering my input, opinions, and ideas.						✓	✓	✓	
Understanding my role as a team member, and contributing to team success.							✓	✓	
Asking for help when I need it.								✓	
Identifying strategies to resolve conflicts with my coworkers.							✓	✓	
Understanding my responsibilities and being accountable for my work.							✓	✓	

SRDC's project – Pay for Success

Partnering with employment service providers to help job seekers gain and maintain employment, Pay for Success followed a program model that rewarded providers not only for employment outcomes, but also for helping job seekers reach a series of intermediate in-program success indicators (or milestones) believed to be associated with progress toward sustainable employment. These milestones represented key transition points along the pathway to employment, including early ones most suitable to those fairly distant from the labour market such as creation of an Employment and Learning plan, demonstration of gains in career and job pathfinding skills, demonstration of gains in attitudes toward learning and general wellbeing. At the same time, the model provided strong incentives to adequately prepare and support job seekers to achieve sustainable employment by including rewards for longer-term employment and retention outcomes, including placement in employment in targeted industries, and job retention at three, six, and 12-months. Pay for Success also incorporated a dual customer approach, aligning training with job seeker needs while also preparing them to meet the needs of employers in specific sectors.

This model was adopted by multiple service providers serving different populations with varying degrees of distance from the labour market. One of the providers served a broad group of job seekers, including those close to the labour market requiring minimal assistance, as well as those who required more intensive assistance. Reflecting the range of job seeker needs, this service provider offered a broad continuum of employment and training services, with Essential Skills assessments done at two different points along the training pathway: 1) a set of generic standardized assessments to document pre-employment classroom training gains; and 2) an employer-rated, industry-contextualized tool that embedded assessments of eight of the nine Essential Skills within a set of sector-specific job tasks to document progress during workplace technical training.

Both assessments represented key performance milestones, with the second one in particular serving to identify participants' readiness to meet employer–defined performance benchmarks as they transitioned from classroom-based training to work placements. Employer trainers observed and evaluated each participant on a scale from 1 to 5 for each of fourteen job performance areas (illustrated in Table 11, along with the Essential Skills underlying each performance area). At the end of the assessment, the examiner summed up the total score and followed this evaluation scale to interpret each participant's total score:

- 0-14: Candidate does not demonstrate necessary skills

- 15-28: Candidate demonstrates a minimal amount of skill in most of the required areas

- 29-42: Candidate demonstrates progress, but still requires further development

- 43-56: Candidate is making progress, but need to work on consistency

- 57-70: Candidate has demonstrated desired occupational specific essential skills.

Table 11   Occupation-specific Essential Skills **–** Sewing Machine Operator

	Reading	Writing	Document use	Numeracy	Digital Literacy	Oral Communication	Thinking Skills	Working with Others	Continuous Learning
Reads and understands posted memos in training area, and performs the directions given.	✓								
Writes reminders to themselves about the sewing machine operation and stitches they use.		✓							
Reads and understand bundle labels.			✓						
Measures seams to ensure acceptable allowances are achieved.				✓					
Demonstrates they are able to manually monitor their own performance.				✓					
Recognize common angles and follow directions given by trainer to ensure desired angles are achieved.				✓					
Estimate the amount of time required to complete the tasks given at their station to ensure they spend little to no off-standard time waiting for work.				✓					
Asks clarifying questions when learning a new operation.						✓			
Requests feedback or instruction as required and can restate what they heard.						✓			
Shares resources with other learners.								✓	
Is respectful and cooperative with teammates and trainers.								✓	
Responds positively to the technical training process and demonstrates interest in the training sessions.									✓
Remember sequence of steps for processes as demonstrated by their trainer.							✓		
Is safety conscious and can apply company policy and rules to his or her own work practices.							✓		

Another service provider adapted the Pay for Success model to align with the needs of clients with complex and severe barriers to employment such as those on long-term income assistance, and individuals with very low levels of education and substantial skills gaps. The model was adapted to create a 'bridging' pathway enabling participants to build foundational skills to transition from life stabilization programs into job search programs.

This service provider integrated Essential Skills training with other employability skills training, following a pathway that focused primarily on enhancing participants' readiness to engage in further learning. Milestones on this pathway captured gains in not only Essential Skills (assessed

using objective, standardized measures) but also in psychosocial attributes such as career and job pathfinding self-efficacy, receptivity to continuous learning, and general wellbeing. Unlike the first provider which targeted sustained employment in a specific sector, this service provider did not have an ultimate employment goal for participants, and so considered it sufficient to use a single set of generic, standardized Essential Skills assessments to document preparation for further learning.

Augmenting the ES framework with assessments of psychosocial attributes related to work readiness

Researchers, educators, policy-makers, and practitioners who are interested in predicting academic and work outcomes have traditionally focused on the core essential skills, such as literacy and numeracy, and on fundamental cognitive skills, such as those measured by IQ and achievements tests. While these skills undoubtedly contribute to school and work success, the last few decades have seen a shift towards psychosocial skills and understanding their role in socioeconomic outcomes.

Psychosocial skills underlie the soft essential skills of oral communication, thinking, working with others, and continuous learning. However, it extends beyond these competencies to include characteristics not clearly represented in the ES framework, such as being self-disciplined, responsible, positive, flexible, and adaptable. Psychosocial skills can be just as important as cognitive skills and they are often valued more so by employers. These kinds of skills and attributes are associated with better job performance, higher wages, and greater employability. They are considered critical for academic and job success in the technological and knowledge economy of the 21st century. Although we have adopted the term psychosocial skills, they are also commonly referred to as soft skills, personality traits, non-cognitive abilities, character skills, personal qualities, intra- and inter-personal skills, and socioemotional skills.

The dominant framework used in the organizational, economic, and personality literature is the Big Five Model (Goldberg, 1971) or the Five Factor Model (McCrae & Costa, 1996)[7]. The five dimensions, forming the acronym *OCEAN*, are listed and described briefly below in Table 12 (APA Dictionary of Psychology, 2018). Each of these dimensions are made up of a number of lower-order facets. For example, conscientiousness can include the lower order facets of self-efficacy, orderliness, dutifulness, achievement striving, self-discipline, and cautiousness. The number of facets and the lower level constructs they represent can vary across personality models and inventories. Some examples are provided in the table below.

---

[7]   These terms are often used interchangeably, but there is a distinction. The Big Five Model views the dimensions as descriptions of behaviour and as a taxonomy of individual differences, whereas the Five Factor Model sees these dimensions as psychological entities with causal power.

Table 12   Big Five/Five Factor Model

Domain	Facets	Description
Openness to experience (Intellect)	Imagination, artistic interests, emotionality, adventurousness, intellect	The tendency to be open to new aesthetic, cultural, or intellectual experiences
Conscientiousness	Self-efficacy, orderliness, dutifulness, achievement-striving, self-discipline, cautiousness	The tendency to be organized, responsible, and hardworking
Extraversion	Friendliness, gregariousness, assertiveness, activity level, excitement-seeking, cheerfulness	Orientation of one's interests and energies toward the outer world of people and things rather than the inner world of subjective experience; characterized by being relatively outgoing, gregarious, sociable, and openly expressive
Agreeableness	Trust, morality, altruism, cooperation, modesty, sympathy	The tendency to act in a cooperative, unselfish manner
Neuroticism (Emotional stability)	Anxiety, anger, depression, self-consciousness, immoderation, vulnerability	Characterized by chronic levels of emotional instability and proneness to psychological distress; emotional stability is the predictability and consistency in emotional reactions, with absence of rapid mood changes

Source: Description of big five from APA Dictionary of Psychology (2018) and facets from IPIP-NEO-PI-R (https://ipip.ori.org/newNEOKey.htm)

Some researchers (e.g., Costa & McCrae, 1992; Goldberg, 1993; John, 1990) argue that the big five are the "longitude and latitude of personality traits, by which all more narrowly defined traits may be categorized" (Almlund, Duckworth, Heckman, & Kautz, 2011, p. 18). That is, these five broad dimensions can subsume all lower-level traits, even those not directly represented as a facet. For example, self-esteem and locus of control are commonly considered to be aspects of emotional stability and are highly correlated with this dimension. However, there are other characteristics (e.g., learning goal orientation) that might not fit in as easily within the framework.

The big five framework is useful not only for organizing lower-level traits, but for conceptualizing mid to high level work competencies, such as customer service or teamwork. Combinations of facets, cutting across domains, can be mapped onto academic and work competencies, an approach used to develop work-related assessments. For example, competency in customer service is likely built on facets of altruism, cooperation, and friendliness (facets of agreeableness and extraversion). The big five framework covers most of the range of personality and can easily encompass most work competencies. This approach can be applied to soft essential skills. For example, Table 13 shows a preliminary mapping of personality facets to each of the soft essential skills, based on content and on face validity.

Table 13   Example of facets underlying soft essential skills

Essential skill	Skill description	Domains and facets	Personality description
Oral communication	The skills needed to exchange thoughts and information with other people by speaking, listening and using non-verbal cues, such as body language	Conscientiousness (self-efficacy)	Handling tasks smoothly, coming up with good solutions, and having good judgement and understanding of consequences and situations.
		Extraversion (friendliness, assertiveness,)	Feeling comfortable around people, interested in others, and warms up quickly. Having things to say, taking charge, and influencing others.
		Agreeableness (cooperation, altruism)	Being easily satisfied, not pushy or confrontational, and not insulting others. Anticipating the needs of others, making people feel comfortable, and taking time for them.
Thinking	The skills needed to solve problems, make decisions, think critically, plan, remember details, and find information	Openness (intellect)	Interest in challenging material, thinking, solving complex problems, and taking in information.
		Conscientious (self-efficacy, cautiousness, orderliness)	Getting things done, coming up with good solutions, and having good judgement and understanding of consequences and situations. Avoiding mistakes, and thinking before acting. Being orderly, and doing things according to a plan.
Working with others	The skills needed to interact with other people (one or more)	Extraversion (friendliness, cheerfulness, assertiveness)	Feeling comfortable around people, and interested in being with others. Being positive about life and joyful. Having things to say, taking charge, and influencing others.
		Agreeableness (trust, altruism, cooperation)	Trusting others and believing in good intentions. Helping others, anticipating their needs, and making them feel welcome. Being easily satisfied, not pushy, and not insulting others.

Essential skill	Skill description	Domains and facets	Personality description
Continuous learning	The skills needed to continually develop and **improve one's skills and** knowledge in order to work effectively and adapt to changes	Openness (intellect, adventurousness)	Interest in challenging material, thinking, solving complex problems, and taking in information. Enjoying changes and trying many new things.
		Conscientiousness (achievement-striving, self-efficacy)	Working hard, turning plans into actions, and being motivated to succeed. Getting things done, exceling at tasks, having good judgement and understanding of consequences and situations.

Source: Descriptions of soft essential skills from https://www.canada.ca/en/employment-social-development/programs/essential-skills/profiles/guide.html

Although emotional stability is not represented above, it likely affects overall work competence as it includes qualities such as being able to adapt to new situations, coping well, staying calm, readily overcoming setbacks, and being comfortable in many situations and with oneself. A more informed and valid mapping of these skills can be achieved using systematic theory- and data-driven methods. This could be a valuable exercise that can inform how soft essential skills could be measured using psychosocial and personality instruments, providing an alternative or supplement to traditional measures used within the ES framework. The advantage of psychosocial measures is that many have good psychometric properties, are well-validated, and have a documented history of use in the research and validation literature.

The evidence for psychosocial attributes

The evidence for the value and power of psychosocial skills comes from multiple sources. First, from the perspective of the employer, psychosocial skills are often rated as more important in hiring, retention, and satisfaction with employees than other types of skills. Kautz, Heckman, Diris, ter Weel, & Borghans (2014) provide a brief review of these employer surveys. For example, employers in the United States reported that attitude, communication skills, responsibility, integrity, and self-management were more important than basic skills, grades, or test scores (Holzer, 1997; Zemsky, 1997). Psychosocial skills might be especially important for entry level and hourly positions. According to one survey, 69% of employers rejected hourly applicants because they were lacking skills such as showing up on time and work ethic. This was double the percentage of employers who rejected applicants because of inadequate reading and writing skills (Barton, 2006). These findings are similar to those in the United Kingdom, where skill gaps are reported mostly in communication, teamwork, and problem-solving skills, rather than numeracy and literacy (Hillage, Regan, Dickson, & McLouglin, 2002; Westwood, 2004). Furthermore, the top ten most

important work competencies[8] based on ratings in the Occupational Information Network (O*NET) database, a job analysis operated and maintained by the U. S. Department of Labor, are all psychosocial skills (Burrus, Jackson, Xi, & Steinberg, 2013). This includes dependability, attention to detail, integrity, cooperation, initiative, self-control, stress tolerance, analytical thinking, adaptability/flexibility, and persistence. These psychosocial skills, called work styles in O*NET, had higher importance ratings from incumbents and analysts and were more universally valued across occupations than all other work abilities, skills, and knowledge.

Some of these psychosocial skills are part of what have been called 21st century skills. From the early 1980s, governmental, academic, non-profit, and corporate organizations have worked to identify a set of skills and competencies needed for academic, work, and life success in the 21st century. The movement began in the United States (e.g., Partnership for 21st Century Learning), but has spread across other countries, including Canada (e.g., Canadians for 21st Century Learning), and internationally (e.g., Assessment and Teaching of 21st-Century Skills.). When the various Canadian, American, and international frameworks are compared, there are several key skills that are represented across almost all frameworks. In an analysis undertaken by the Ontario Ministry of Education (2016), comparing 18 Canadian frameworks and 7 international frameworks, the top four skills endorsed by frameworks were: collaboration and teamwork, communication, creativity/innovation, and critical thinking. In another comparison by Hanover Research (2011) of 6 mostly American frameworks, the top four skills were: collaboration and teamwork, creativity and imagination, critical thinking, and problem solving. Finally, using an empirical approach with O*NET ratings (i.e., principal components analysis of importance ratings), Burrus and colleagues (2013) identified the top 5 skills as: problem solving, fluid intelligence, teamwork, achievement/innovation, and communication. Across frameworks and methodologies, the key skills for success in the future appear to be primarily psychosocial skills and critical thinking or problem-solving skills.

Evidence for psychosocial skills is also supported by the personality and organizational research literature. Since the 1990s, a large number of meta-analyses have demonstrated the ability of personality measures to predict academic and work outcomes (e.g., Barrick & Mount, 1991; Ones, Viswesvaran, & Schmidt, 1993; Salgado & Táuriz, 2014; Trapmann, Hell, Hirn, & Schuler, 2007). Conscientiousness and emotional stability are the most robust predictors of educational and work outcomes (for a brief review, see Almlund et al., 2011). Conscientiousness predicts educational attainment and achievement, while both conscientiousness and emotional stability predict job performance and wages. Meanwhile, other traits, such as openness to experience and agreeableness, predict more specific outcomes, such as attendance at school, or selection into particular careers. In general, personality predicts across occupation and education levels, whereas cognitive ability is more important for jobs that are complex and require higher levels of education. Personality affects job outcomes through educational attainment. For example, research shows that adolescent conscientiousness predicts employment outcomes in adulthood, but academic motivation and educational attainment only accounted for less than 9% of this association (Egan,

---

[8] Ratings were analyzed for Zone 3 to 5 jobs, those requiring some post-secondary education.

Daly, Delaney, Boyce, & Wood, 2017). Personality likely affects work outcomes additionally through occupational matching, absenteeism, turnover, and job search. A consideration to bear in mind is that education itself can affect personality. Education can foster psychosocial skills, which then increases schooling and leads to enhanced labour market outcomes. Cognitive and psychosocial skills tend to go hand in hand, such that those who have high cognitive skills and do well in school also score highly on psychosocial skills (Todd & Zhang, 2018).

One weakness of this body of research is that most of the evidence are associations between job outcome and psychosocial skills. It cannot speak to causality, which is important in policy decisions involving the evaluating and funding of social and educational programs (Kautz et al., 2014). We can, however, look to research on cognitive and psychosocial interventions, where programs effects on psychosocial skills have led to improved life outcomes. In a review of early childhood, elementary school, adolescent, and young adult interventions, Kautz et al. (2014) concluded that early interventions tend to work more through psychosocial skills than cognitive skills. For example, the Perry Preschool Program was a comprehensive socioemotional intervention for disadvantaged preschoolers during the 1960s. The initial evidence showed significant gains in IQ in the experimental group compared to the control group, but these effects gradually dissipated and by age 10, the two groups were no longer different. However, there were other lasting psychosocial gains (e.g., fewer externalizing behaviours, greater academic motivation) and when looking at long-term outcomes, the experimental group was more successful than the control group (Heckman, Malofeeva, Pinto, & Savelyev, 2010). Another example comes from the General Educational Development (GED) testing program in the United States. Based on test scores, GED recipients are as smart as traditional high school graduates, but even after adjusting for cognitive skill, GED recipients perform much more poorly in the labour market and other domains of life (Heckman, Humphries, & Kautz, 2014c). GED recipients are much more similar in psychosocial skills to dropouts, and it is deficits in skills such as persistence, organization, and showing up on time, that lead them to underperform later in life. Based on their review of intervention programs from early childhood through young adulthood, Kautz and colleagues (2014) conclude that psychosocial skills are more malleable in later stages of life than cognitive skills, and a more effective target for intervention programs.

Methodological considerations for psychosocial instruments

Psychosocial measures are unlike those used for cognitive tests and even core essential skills. Most are subjective self-report measures, typically asking people to rate themselves on a variety of attitudes, skills, behaviours, or personality statements. However, when people rate themselves on different attributes, they do so by comparing themselves to other people. For example, people have to interpret what "impulsive" means, and this can differ depending on people's reference point and what is common in the people around them. This can lead to *reference bias*, which makes it difficult to compare results across groups of people. Another concern with self-report measures is *faking*, where people identify the desirable traits or skills and endorse those at higher rates. Some researchers believe faking is a real concern (e.g., Morgeson et al., 2007). In experimental contexts, participants instructed to fake doing well can raise their scores (Viswesvaran & Ones, 1999). Meanwhile, others argue that in real settings, people do not fake as much as we think (Hogan,

Barrett, & Hogan, 2007). For example, researchers showed that real job applicants only improved their scores by about 5% or less when they completed a personality measure for the second time, after being rejected for the job the first time 6 months ago. Ultimately, faking is a concern primarily for personnel selection and other high-stakes contexts. Faking is less of a concern when psychosocial assessments are used for training, skill development, and program evaluation.

There are, however, various methods that can be used mitigate the effects of faking. One method is to administer social desirability scales along with a psychosocial assessment to either correct for faking (Ellingson, Sackett, & Hough, 1999), which does not appear very effective, or to use high social desirability scores to identify tests that might not be valid (NCS Pearson, Inc., 2007a). Others use low correlations between a psychosocial assessment and social desirability scales to demonstrate low probability of faking. Another method is to include warnings against faking as part of assessment instructions (McFarland, 2003). Finally, researchers are developing fake-resistant tests, such as using forced-choice format. In many personality tests, individuals are given a series of statements or adjectives and asked to rate themselves using a 5- or 7-point scale on how strongly they agree or disagree with that description of themselves. In forced-choice formats, individuals are given two or more statements that are equally socially desirable and asked to endorse the one most like or least like themselves. Evidence suggests that these forced-choice assessments are valid and can predict job outcomes (see Salgado & Táuriz, 2014 for meta-analytic review).

Another common debate in the field of psychosocial and personality assessment is the bandwidth-fidelity dilemma (Cronbach & Gleser, 1957). When you use broad measures (e.g., the big five domains), they cover the majority of variation in personality and can predict broad criteria (e.g., job performance) with moderate validity. However, for maximum validity, there needs to be greater fidelity between the measure and criterion. This can be found in narrow measures (e.g., facets of the big five domains), which assess less variation in personality, but assess it more precisely. There are several views regarding this trade-off (see Salgado et al., 2014). One view is that broad factors can predict both broad (e.g., overall job performance) and narrow (e.g., counterproductive work behaviours) performance criteria (e.g., Ones & Viswesvaran, 1996). A second view is that narrow measures predict narrow criteria better than broad measures, and have incremental predictive power over broad measures (e.g., Ashton, 1998; Paunonen, Rothstein, & Jackson, 1999). Finally, a third view is that predictor measures and performance criteria should be aligned, such that that broad measures should be used to predict broad outcomes, and narrow measures to predict narrow outcomes (e.g., Hogan & Roberts, 1996). There is no clear answer to this debate, and the choice of broad or narrow measures will likely depend on the purpose of the assessment, the performance outcomes of interest, and the availability of measures and resources. Reviewing previous research and field use can be helpful to ensure the measures chosen have been successfully used in the past in similar contexts.

Selecting psychosocial measures

There are several types of psychosocial measures that organizations can use in work readiness and job performance contexts. The type of measure used, whether ready-made or customized, is determined by the intended use of the measure (e.g., academic and research, evaluating programs, identifying talent), the stakes involved, and the technical and financial resources available. The

main types of psychosocial measures identified in our review include personality inventories and other research instruments, work readiness assessments, and talent assessments. These are briefly described below. In the subsequent sections, several examples of each type of measure are described.

Personality inventories and other research instruments

One approach is to use existing validated personality inventories that are available either commercially or through the public domain. Personality inventories cover a wide range of characteristics and usually measure both broad and narrow constructs (e.g., the big five domains and their corresponding facets). The big five domains are useful for assessing general psychosocial competency whereas measurement at the facet level is useful for a detailed assessment. Organizations can also measure specific domains or facets of interest. These inventories generally offer good psychometric properties when the scales are administered as intended.

Many personality inventories are proprietary and require licensed training to administer, but there are a number of well-known and commonly-used measures that are publicly available (e.g., Big Five Inventory). In particular, the International Personality Item Pool (IPIP) is a large rich database of personality statements or items that can be used to create personality scales. It also provides access to free analogs of well-validated proprietary personality inventories (for more information, see below). In the academic literature, there is also a large number of research instruments that measure lower-level constructs (e.g., self-esteem, grit). Some are considered to fall within the big five framework and others are not. Many of these instruments are accessible online or through access to journals. They also generally have good psychometric properties, with published studies of their validation, and history of use by other researchers. Some instruments were developed to measure work-specific constructs, such as occupational self-esteem or work ethic. Others measure general constructs that have been applied in work contexts, such as self-control or grit. If possible, organizations should select measures that have been used in workforce samples and that have been associated with job outcomes. These instruments can be useful for organizations who are interested in measuring specific characteristics rather than broad assessments of abilities.

Personality measures, especially those related to the Big Five framework, are increasingly being used to complement essential skill measures in the context of international assessment initiatives such as PIAAC. As discussed in greater detail below, self-reported personality can perform as well as essential skill scores in predicting a variety of life outcomes.

Work-readiness assessments

Another approach is to use existing work readiness or employability assessments, which measure general skills and attributes that are considered important for success in the labour market. They are often based on needs identified by employers, the experiences and insights of professionals who work in employment services and skill training programs, and what has been shown to be related to employability. These assessments are useful for students who are on the verge of graduating and entering the job market for the first time, and for adults who are re-entering the job market after training or unemployment. They can be used by career professionals providing career counselling

services or skills training to clients, and by program administrators to assess the effectiveness of their services. Although these assessments can include core essential skills, such as literacy and numeracy, many emphasize psychosocial work readiness, and qualities that contribute to strong work ethic, professionalism, and customer service. Given the social and labour market contexts, these assessments have been often been developed with the involvement of governmental departments and non-profit organizations. Existing work readiness assessments likely involve a moderate cost.

Work readiness assessments are driven by the identified skills and characteristics demanded by industry and related to job success. Organizations interested in creating their own work readiness assessment can gather this information through survey or interview with stakeholders in the industry of interest, and career and training professionals. Another source of information is both the academic and grey literature, and existing work readiness assessments. Once key characteristics for success have been identified, organizations can then create, select, or adapt measures of those skills, such as by using the IPIP or other research instruments. If customized measures are used, the validity of these measures will need to be tested during use in the field.

Talent assessments

A third approach is to use talent assessments, which are proprietary commercial tests designed for organizations and companies interested in using personality and other psychosocial factors to guide personnel selection. This can include assessing the fit of candidates to different job profiles. These assessments, unlike work-readiness assessments, are high-stakes, with the results being directly used to inform hiring and other decisions that affect individual test takers. There are several large testing or psychometric companies who have developed talent assessments or administer other validated assessments, including ETS, ACT, and Pearson. A characteristic of these talent assessments by large psychometric companies is that the development and testing of these assessments tend to be more psychometrically rigorous and comprehensive. Personality characteristics and taxonomies are mapped to work domains and tasks, and validated with workforce and organizational samples. However, these assessments have a high cost associated with them and are more suited for situations where precision is needed in predicting outcomes or fit for specific occupations or industry.

Organizations that are interested in developing their own measures informed by both personality and job performance models might first begin by mapping the jobs profiles of the occupation or industry of interest (e.g., using the National Occupation Classification system or stakeholder interviews) and specifying specific work components required for successful job performance. Following ETS's formative construct approach (Naemi, Burrus, Kyllonen, & Roberts, 2012), lower-level facets can be identified that composited to create formative mid-level constructs that map onto educational or workforce outcomes. In the example below (Figure 23), the facets of social dominance, organization, and concern/respect are composited to form the construct of leadership, which predicts both achievement and attitudes.

Figure 23 **An example of ETS's formative construct approach**
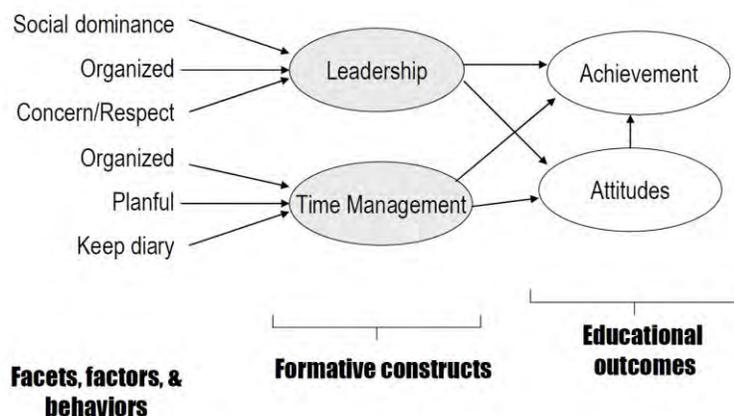
Source: p. 2 of Naemi, Burrus, Kyllonen, & Roberts (2012).

Although there is limited information about the development of these talent assessments, sometimes the conceptual structure is shared. For example, one of ETS's talent indices is *initiative and perseverance*, which is a composite of the facets diligence, assertiveness, and dependability. Examining the model structures of these assessments can help guide an organization's own development of a customized measure. Some development methods may require more technical psychometric expertise, but the necessity of this rigor will depend on the purpose of the assessment.

In the following sections below, several examples of personality inventories, work readiness assessments, and talent assessments are briefly reviewed. This information can be used to understand the broad range of assessments, provide sources for example assessments, and guide an organization's own process of developing a customized measure.

Examples of personality inventories and other research instruments

Personality inventories

There are many personality inventories that have been developed. Some are proprietary and commercial, with significant costs attached to them, while others are freely available in the public domain. Personality assessments also vary in length. For example, one of the most well-known personality inventories based on the big five model is the NEO Personality Inventory-Revised (NEO PI-R) with 240 items. In contrast, the Ten Item Personality Inventory (TIPI) has only 10 items. However, the NEO PI-R also measures 30 facets, whereas the TIPI only measures the big five domains. Further, because of the small number of items in the TIPI, the instrument is designed not to maximize reliability or factor structure, but content and criterion validity. The choice of test length will depend on the needs and priorities of the organization or project. Another factor is whether an organization is interested in the broad dimensions of the big five model, or in specific facets. Table 14 provides a brief description of some of the more commonly used personality

inventories that are based on or builds on the big five model. However, there are many other well-known inventories including the California Psychological Inventory and the Minnesota Multiphasic Personality Inventory.

Table 14   Examples of big five related personality inventories

Instrument	Description	Structure
NEO Personality Inventory-Revised (NEO PI-R) *Paul Costa & Robert McCrae*	Commercial 240 items, 3 validity items, 5-point rating scale Takes 30-40 minutes to complete Pencil and paper Many versions available (e.g., NEO-PI-3, NEO-FFI-3), including NEO-PI-3: 4FV, which focuses on four factors, excluding neuroticism, for use in employment, career, and training contexts	5 domains and 30 facets: ▪ Conscientiousness (competence, order, dutifulness, achievement striving, self-discipline, deliberation) ▪ Openness to Experience (fantasy, aesthetics, feelings, actions, ideas, values) ▪ Agreeableness (trust, straightforwardness, altruism, compliance, modesty, tender-mindedness) ▪ Extraversion (warmth, gregariousness, assertiveness, activity, excitement-seeking, positive emotions) ▪ Neuroticism (anxiety, angry hostility, depression, self-consciousness, impulsivity, vulnerability)
Big Five Inventory (BFI, BFI2) *Oliver John*	Free online (https://www.ocf.berkeley.edu/~johnlab/bfi.htm) 44-item, 10-item, and new 60-item versions available, 5-point rating scale	5 factors: ▪ Conscientiousness ▪ Agreeableness ▪ Extraversion ▪ Openness to Experience ▪ Neuroticism
Big Five Aspect Scales (BFAS) *Colin DeYoung, Lena Quilty, & Jordan Peterson*	Free online through IPIP (https://ipip.ori.org/BFASKeys.htm) An online administrated version with generated reports available for a fee on https://understandmyself.com 100-items, 5-point rating scale	5 domains with two aspects each: ▪ Conscientiousness (industriousness, orderliness) ▪ Openness (intellect, openness) ▪ Extraversion (enthusiasm, assertiveness) ▪ Agreeableness (compassion, politeness) ▪ Emotional stability (volatility, withdrawal)

Instrument	Description	Structure
Big Five Mini-Markers *Gerard Saucier*	Free online through third party sites; available in referenced journal article 40 items (trait-descriptive adjectives), 9-point rating scale **Short version of Lew Goldberg's original** 100-trait descriptive adjectives	5 factors: ▪ Conscientiousness ▪ Agreeableness ▪ Extraversion ▪ Intellect or Openness ▪ Emotional Stability
Ten Item Personality Inventory *Sam Gosling, Jason Rentfrow, & Bill Swann*	Free online (https://gosling.psy.utexas.edu/scales-weve-developed/ten-item-personality-measure-tipi/) 10-items, 7-point scale Some norms available online Short instrument designed to optimize validity (especially content and criterion validity), and given small number of items, likely will have low alpha reliability and poor factor structure	5 domains (2 items each): ▪ Conscientiousness ▪ Agreeableness ▪ Openness to experience ▪ Emotional Stability ▪ Extraversion
HEXACO-PI-R *Kibeom Lee & Michael Ashton*	Free for non-profit academic use (may contact authors for non-academic use) (http://hexaco.org/) 200-item, 100-item and 60-item versions for self- and other-report 5-point rating scale	6 dimensions, 25 facets: ▪ Honest-Humility (sincerity, fairness, greed avoidance, modesty) ▪ Emotionality (fearfulness, anxiety, dependence, sentimentality) ▪ Extraversion (social self-esteem, social boldness, sociability, liveliness) ▪ Agreeableness (forgivingness, gentleness, flexibility, patience) ▪ Conscientiousness (organization, diligence, perfectionism, prudence) ▪ Openness to Experience (aesthetic appreciation, inquisitives, creativity, unconventionality) ▪ Additional facet: Altruism (versus antagonism)

International Personality Item Pool (IPIP)

Because many of the most popular personality inventories are proprietary, the IPIP
(https://ipip.ori.org/, see Figure 24 for an illustration of its interface) is a valuable tool for
organizations who do not have the resources for commercial tests, or who want to create
customized instruments. The IPIP is the result of an international effort to provide a resource that is
in the public domain and openly available for scientific and commercial purposes. This access
enables continued development and refinement of personality inventories and assessments by the
research community. A detailed history and rationale of the project is available online. Today the
IPIP has over 3, 000 items, and over 600 studies have been published using these items. One of the
most powerful features of the IPIP is the construction of analog scales of popular, often proprietary,
personality inventories. These analog versions of proprietary inventories were developed by
correlating IPIP items with those from personality inventories. Items were selected based on the
strength of their correlations with the various scales, avoiding redundancy of items, content
analysis, and reliability analysis. The IPIP has analog versions of the NEO-PI-R, BFAS, and HEXACO-
PI mentioned above. In total there are currently 463 IPIP analog scales and 26 personality
inventories represented. There are 274 traits or facets that have at least one analog scale.
Examining the list of traits provides an easy way of finding items for a particular construct, and
identifying which personality inventories measure the construct.

The IPIP provides a lot of flexibility and since these items are in the public domain, individuals are
free to alter wording, choosing rating scales of their choice, and create their own scales or adapt
existing ones. However, the further one strays from the pre-constructed scales, the less one can rely
on the validity and reliability information provided by IPIP and by proxy, the original personality
inventories.

Figure 24   **IPIP's homepage and content available**

Source: https://ipip.ori.org/

Other research instruments

The personality, psychology, and organizational literature offers a large number of instruments designed to measure all types of attitudes, traits, behaviours, and skills. These are generally lower-level traits, rather than high level domains, such as the big five. The benefit of using these instruments is that, once access is gained to the original research, they are usually free to use. Some of the more widely-used and popular measures are freely available on the Internet. Another advantage is that organizations can be informed by and build on prior research using these measures. Some measures have been translated into multiple languages and used world-wide across many contexts. However, most of this information is likely only accessible through academic journals and perhaps some grey literature. This information can be valuable when deciding which measures to use for specific populations or outcomes of interest, and to compare and contextualize results.

Table 15 provides examples of some common research instruments that are used, focusing on constructs related to conscientiousness and emotional stability, which show the most robust relations with academic and work outcomes.

Table 15   Examples of other research instruments

Research instrument	Big five	Description	Structure and sample items
Grit			
Grit Scale *Duckworth & Quinn (2009)*	Conscientiousness	Available free online (https://angeladuckworth.com/grit-scale/) 8 items, 5-point rating scale Earlier version exists with 12 items	2 facets: ▪ Perseverance of effort ▪ Consistency of interest Sample items: ▪ **Setbacks don't discourage me. I don't** give up easily (perseverance) ▪ I often set a goal but later choose to pursue a different one (consistency)
Self-Control			
Self-Control Scale (SCS); Brief Self-Control Scale (BSCS) *Tangney, Baumeister, & Boone (2004)*	Conscientiousness	Available in referenced journal article 36 items, 5-point rating scale 13-item brief version also available	Sample items: ▪ I am good at resisting temptation ▪ **I can't carried away by my feelings**

Research instrument	Big five	Description	Structure and sample items
Work Ethic			
Multidimensional Work Ethics Profile (MWEP); MWEP-SF (Short form) *Miller, Woehr, & Hudspeth (2001)* *Meriac, Woehr, Gorman, & Thomas (2013)*	Conscientiousness	Available in referenced journal articles 65 items, 5-point rating scale 28-item version available	7 dimensions: ▪ Self-reliance ▪ Morality/ethics ▪ Leisure ▪ Hard work ▪ Work centrality ▪ Wasted time ▪ Delay of gratification Sample items: ▪ Any problem can be overcome with hard work (hard work) ▪ A distant reward is usually more satisfying than an immediate one (delay of gratification)
Self-Efficacy			
General Self-Efficacy Scale (GSE) *Schwarzer & Jerusalem (1995)*	Emotional stability	Available free online (http://userpage.fu-berlin.de/~health/selfscal.htm) 10 items, 4-point rating scale	Sample items: ▪ I can always manage to solve difficult problems if I try hard enough ▪ I am confidence that I could deal efficiently with unexpected events
General Self-Efficacy Scale (SGSE) *Sherer et al. (1982)*	Emotional stability	Available in referenced journal article 17 items, typically 5-point rating scale (developed using 14-point rating scale)	Sample items: ▪ When I make plans, I am certain I can make them work ▪ When unexpected problems occur, I **don't handle them well**
New General Self-Efficacy Scale (NGSE) *Chen, Gully, & Eden (2001)*	Emotional stability	Available in referenced journal article 8 items, 5-point rating scale	Sample items: ▪ When facing difficult tasks, I am certain that I will accomplish them ▪ I will be able to successfully overcome many challenges

Research instrument	Big five	Description	Structure and sample items
Occupational Self-Efficacy Scale (OCCSEF) *Schyns & Collani (2002)*	Emotional stability	Available in referenced journal article 19 items, 6-point rating scale An 8-item and 6-item version also available Based on items from both GSE, SGSE, and two other scales; adapted to workplace	Sample items: ▪ When I set goals for myself in my job I rarely achieve them ▪ No matter what comes my way in my **job, I'm usually able to handle it**
Self-Esteem			
Rosenberg Self-Esteem Scale (RSE) *Rosenberg (1965)*	Emotional stability	Available in referenced journal article (but can be found online on third party sites) 10 items, 4-point rating scale	Sample items: ▪ On the whole, I am satisfied with myself ▪ I certainly feel useless at times
Single-Item Self-Esteem Scale (SISE) *Robins, Hendin, & Trzesniewski (2001)*	Emotional stability	Available in referenced journal article (but can be found online on third party sites) 1 item; a proxy for RSE	Sample item: ▪ I have high self-esteem
Locus of Control			
Rotter Locus of Control Scale *Rotter (1966)*	Emotional stability	Available in referenced journal article 23 sets of items (plus 6 filler items), forced-choice pairs of items	Sample pair of items: ▪ In the long run, people get the respect they deserve in this world ▪ **Unfortunately, an individual's worth** often passes unrecognized no matter how hard he tries
Levenson IPC Scale *Levenson (1973)*	Emotional stability	Available in referenced journal article 24 items, 6-point rating scale	3 dimensions: ▪ Internality ▪ Powerful others ▪ Chance Sample items: ▪ I can pretty much determine what will happen in my life (internal) ▪ I feel like what happens in my life is mostly determined by powerful people (powerful others)

Research instrument	Big five	Description	Structure and sample items
Work Locus of Control Scale (WLCS) *Spector (1988)*	Emotional stability	Available in referenced journal article 16 items, 6-point rating scale	Sample items: ▪ Getting the job you want is mostly a matter of luck ▪ Promotions are given to employees who perform well on the job
Core Self Evaluations			
Core Self-Evaluation Scale *Judge, Erez, Bono, & Thorensen (2003)*	Emotional stability	Available free online (http://www.timothy-judge.com/CSES.htm) 12 items, 5-point scale	Core self-evaluations encompass four core self-evaluation traits: self-esteem, generalized self-efficacy, neuroticism, and locus of control. However, they are no separate scales for each trait. Sample items: ▪ I am confident I get the success I deserve in life ▪ Sometimes when I fail I feel worthless

Personality measurement in the context of international assessments

The measurement of personality and other non-cognitive skills has become an increasing priority in the context of large-scale international studies. For example, the four countries that have carried out longitudinal follow-ups of their Programme for the International Assessment of Adult Competencies (PIAAC) samples – Germany, Poland, Italy, and Canada – have all included short self-rated personality inventories in their follow-up surveys.

In Canada, PIAAC follow-up consists of three waves of the Longitudinal and International Study of Adults (LISA). The first wave in 2014 collected data on a variety of socioeconomic outcomes, including labour market, education and training, and health, as well as a self-assessed Big Five inventory. The other participating countries have implemented similar measurement frameworks, allowing for exploration of the links between life outcomes and both cognitive skills (in the form of objectively assessed Essential Skills competencies) and non-cognitive personality dimensions in large representative samples of their adult populations.

As of this report, we are unaware of any analysis and dissemination of Canadian data, but results from other countries have been striking. In both the 2014 German PIAAC-L and 2014 Polish POSTPIAAC follow-ups, respondents completed a 15-item short version of the Big Five Inventory (BFI). Results confirmed that differences in personality traits – especially in the Big Five dimensions of conscientiousness and emotional stability – are important predictors of life outcomes. For example, in the Polish sample, after controlling for demographic factors

conscientiousness and emotional stability are both significant and positively correlated with labour market participation, wages, and job satisfaction. Numeracy on the other hand, is only significantly correlated with wages, and even then to a lesser extent than conscientiousness (Palczynska & Swist, 2016). Similarly, in the German sample being employed full-time was positively related to conscientiousness and emotional stability as well as numeracy (Rammstedt, Danner, & Lechner, 2017).

The emerging importance of psychosocial skills in international assessment is reflected in the content of the online version of PIAAC, the Education and Skills Online Assessment (E&S Online). E&S Online is designed to provide individual-level results that are linked to PIAAC measures of cognitive skills, but also includes non-cognitive skills modules that were field tested and validated with representative populations in several countries, including Canada. The E&S Online Behavioral Competencies module provides scores across 13 personality traits related to Big Five domains and thought to be related to success in workplace and education (Table 16).

Table 16   Behavioral Competencies Module, Education and Skills Online Assessment Tool

Broad Personality Category	Lower-level trait	Description
Conscientiousness	Diligence	Diligence describes behaviors associated with working towards objectives. Individuals who are high in diligence tend to be described as hard working, ambitious and confident.
	Organization	Organization describes behaviors associated with maintaining a sense of order as well as an ability to plan work tasks and work activities.
	Dependability	Dependability describes behaviors related to a sense of personal responsibility. Individuals who are high in dependability tend to be reliable and make every effort to keep promises.
	Self-discipline	Self-discipline indicates an ability to be patient, cautious and level-headed. People who are high in self-discipline tend to maintain control at work.
Extraversion	Assertiveness	Assertiveness indicates an ability to take charge at work. People who are assertive are often described as direct, **decisive and "natural leaders."**
	Friendliness	Friendliness indicates an interest in social interactions. People high in friendliness are often interested in meeting new people at work and using this skill for the betterment of the organization.

Broad Personality Category	Lower-level trait	Description
Agreeableness	Generosity	Generosity describes individuals who are willing to offer their time and resources in support of others. People high in generosity tend to be helpful to others at work.
	Collaboration	Collaboration describes individuals who are viewed as trusting and cooperative. People high in collaboration are often easy to get along with and work well on teams.
Emotional Stability	Stability	Stability describes individuals who are relaxed and worry free. People high in stability work well with changing work priorities and manage stress well.
	Optimism	Optimism describes individuals who have a positive outlook and cope well with setbacks. People who are optimistic tend to incorporate feedback well at work.
Openness to Experience	Creativity	Creativity describes behaviors that are inventive and imaginative. People high in creativity tend to be innovators at work.
	Intellectual Orientation	Intellectual Orientation is indicative of an ability to process information and make decisions quickly. People high in intellectual orientation are often viewed as knowledgeable by others.
	Inquisitiveness	Inquisitiveness describes behaviors that relate to being perceptive and curious. People high in inquisitiveness tend to be interested in learning more by attending workshops at work.

Examples of work readiness assessments

Conference Board of Canada

The Conference Board of Canada's Employability Skills 2000+ framework includes skills required to enter, stay in, and progress in today's labour market. Skills are grouped into three domains: fundamental skills, personal management skills, and teamwork skills. Across the three domains, all core and soft essentials skills are represented in varying degrees and forms. For example, teamwork skills includes working with others and participating in projects and tasks. However, the personal management domain adds additional psychosocial skills, specifically demonstrating positive attitudes and behaviours, being responsible, being adaptable, learning continuously, and working safely. The Conference Board developed several self-assessment tools that facilitate skill and career development through self-reflective exercises.

The Employability Skills Toolkit is a suite of practical tools designed for the self-managing learner, or to support the learner-coach relationship for those working with a counsellor or adviser. The toolkit has 9 modules that help individuals reflect on their own interests, values, and skills, identify and assess employability skills, plan and implement skill development activities, and document skill achievements. There are costs associated with the toolkit, but a preview is available online (see Figure 25).

Figure 25   An example from the Employability Skills Toolkit

Employability Skills Toolkit   Module Two—Identify & Reflect

Employability 2000+ Skill	Demonstration of skill
Think & Solve Problems	
• Assess situations and identify problems	*I demonstrate this skill by…* ❏ I haven't yet demonstrated this skill
• Seek different points of view and evaluate them based on facts	*I demonstrate this skill by…* ❏ I haven't yet demonstrated this skill
• Recognize the human, inter-personal, technical, scientific and mathematical dimensions of a problem	*I demonstrate this skill by…* ❏ I haven't yet demonstrated this skill
• Identify the root cause of a problem	*I demonstrate this skill by…* ❏ I haven't yet demonstrated this skill

Source: p. 40 of the Employability Skills Toolkit.

The Skills Credentialing Tool is an online tool designed for students, employees, and job-seekers to assess their own employability skills and create an evidence-based portfolio that can be shared with counsellors or workplace supervisors for skill and career development. The assessment can be accessed online free by anybody with a valid email. For each of 12 groups of skills, individuals rate their proficiency on 5 items using a 5-point scale. There is an opportunity to also provide concrete examples of these skills. A record of the answers are sent to the test-taker by email. However, there is little other support for interpreting results.

Figure 26   An example from the Skills Credentialing Tool

Source: https://survey.conferenceboard.ca/SE/82/sct/

Employability Skills Assessment Tool (ESAT)

The Futureworx society is a non-profit community-based organization that helps diverse clients develop the skills for success in school, work, and life. Futureworx offers employment, educational, and training programs across Nova Scotia (a Nova Scotia Works Employment Services Centre) and works with organizations across Canada and internationally. The tool has been used across all provinces of Canada, as well as several states in the United States.

Funded by the Counselling Foundation of Canada and Réseau pour le développement de l'alphabétisme et des compétences, the Futureworx Society of Truro, Nova Scotia created the Employability Skills Assessment Tool (ESAT) to support the development of nine core soft skills: accountability, adaptability, attitude, confidence, motivation, presentation, stress management,

teamwork, and time management. According to their skills pyramid model, job success is built on essential skills, which are further built on soft skills. It is a cloud-based tool available in English and French designed to be used with clients receiving training to prepare them for employment.

The ESAT tool is adaptable and can be easily adjusted to incorporate different attributes and targeted outcomes, such as performance requirements of specific industries, personality attributes, or the soft Essential Skills including Communication, Working with Others, and Thinking. The tool combines staff assessments, which are observation-based, and learners' self-assessments, which are questionnaire-based, to provide information on learners' behavioural competencies related to employability. Skill assessments between client and staff are compared using radar plots to highlight strengths and areas for improvement from both perspectives (see Figure 27). The tool provides benchmarks aligned with the needs and expectations of most workplaces so that program staff can easily see how ready participants are to participate in the labour market.

The ESAT can be used in multiple program contexts, from those as short as three weeks to create awareness and enhance understanding of employability expectations, to those as long as six months or more to achieve real behavioural changes. In the latter cases, multiple assessments can be used to track progress over time. There is a user-friendly case management interface associated with the tool, giving program staff timely access to participants' results.

Figure 27   An example of reports generated by the ESAT

Employment Readiness Scale (ERS)

The Employment Readiness Scale (ERS) was developed by Valerie G. Ward and colleagues, HRDC, and the BC Ministry of Human Resources. It is designed to help clients identify their strengths and challenges in becoming employed, and to measure changes over time, particularly in the context of evaluating training and other employment programs.

The ERS is based on the Employment Readiness Model that was developed using the international literature (i.e., Canada, United States, Sweden, United Kingdom, Australia), input from subject matter experts, and professionals who provide career and employment services. Employment readiness is determined by four employability factors, five soft skills, and three stress factors (see Table 17).

Table 17   Employment Readiness Model

Employment readiness factors	Description
Employability factors	
Career decision-making	Knowing what type of work suits you
Skills enhancement	Having the skills and education for the work you want
Job search	Having the skills to find work
Ongoing career management	Being able to manage future work life changes
Soft skills	
Self-efficacy	Sense of being able to perform well
Outcome expectancy	Whether or not you expect to succeed and are willing to take responsibility for creating that success
Social supports	Your network and ability to get help
Work history	Your feeling that you have performed well in previous work contexts, paid or unpaid
Job maintenance	Having the skills to keep work once found
Stresses or challenges	
Personal	You can usually address yourself
Environmental	You can manage with help
Systemic	Have to be addressed on a community basis

Source: http://www.employmentreadiness.info/node/3

Of the five soft skills, self-efficacy, and outcome expectancy more directly measure specific psychosocial characteristics. In contrast, social supports, work history, and job maintenance appear to tap into qualities that underlie the ability to maintain a social network, feel you performed well in your work, and maintain a job.

The ERS is an online assessment with separate access portals for clients, agencies and funders, with the ability to generate various reports and action plans. The self-report assessment that clients complete has 45 items that involve ratings on a 5-point scale, and 30 items that involve true and false answers. A short summary of its development and psychometric properties is available online

(Valerie G. Ward Consulting Ltd. & Service-Growth Consultants Inc., 2002), although more information is contained within confidential Crown reports. There are no sample items available online. The full scale, including soft skills, predicts subsequent employment within 12 weeks, and is consistent with staff ratings of soft skills.

Other examples from the United States

There are several other examples of work readiness assessments that were identified in the United States. The two examples described below illustrate how some work readiness assessments can use more objective measures, rather than self-ratings.

The Workplace Readiness Skills (WRS) assessment was developed by the Career and Technical Education Consortium of States (CTECS) and by the Department of Career and Technical Education Services, Virginia Department of Education. Initially developed in 1997, it was updated in 2010, incorporating learnings from the 21st century skills movement, the academic literature, workplace readiness plans in 7 other states, and input from Virginia employers. The assessment was designed for use in high schools, and has been used in the states of Virginia, Nevada, Maine, and Idaho. The assessment consists of 100 multiple choice questions covering three areas: Personal Qualities and People Skills, Professional Knowledge and Skills, and Technology Knowledge and Skills. Whereas most of the core and soft essential skills are primarily represented in the latter two areas, Personal Qualities and People Skills covers additional psychosocial skills: positive work ethic, integrity, teamwork, self-representation, diversity awareness, conflict resolution, creativity and resourcefulness. Unlike many of the other work readiness assessments that are based on self-ratings, the questions are designed with correct and incorrect answers. Some questions present work scenarios, whereas other questions test the individual's understanding of different work values and skills.

Table 18   Sample questions from the WRS

1. Carla believes she is more productive when she listens to music at work, but her boss does not allow it. She should:

   a) try to convince her boss of her point of view.
   b) put on headphones to listen in private.
   c) follow company policy.
   d) listen to it anyway since it is good for productivity.

2. We know that how we communicate is often just as important as what we are trying to say. Avoiding using slang in the workplace is an example of understanding the skill of:

   a) self-representation.
   b) teamwork.
   c) integrity.
   d) diversity awareness.

Source: https://www.ctecs.org/sites/default/files/files/SampleWRS_50.pdf

This assessment was validated by CTECS and the Weldon Cooper Center at the University of Virginia, although this information is not publicly available online. There is, however, conceptual background and implementation material available online.

The Comprehensive Student Assessment System (CASAS) is a nonprofit organization that provides assessments for basic skills and curriculum tools. They have regularly partnered with federal and state government agencies, business and industry, and educational and training institutions. Their Workforce Skills Certification System is designed to assess, profile, develop, and certify work-related academic and soft skills. Academic skills include reading and math skills, and critical thinking and problem solving skills. Soft skills include personal qualities and customer care. Personal qualities includes integrity, responsibility, self-esteem, self-management, and sociability. Customer care includes commitment to quality, customer relations, and decision-making. Soft skills are assessed using a situational judgement test. This involves watching video-based simulations of work scenarios and answering 56 multiple-choice questions in total. An example can be viewed on the assessment developer's website (http://www.learning-resources.com/). This differs also from typical psychosocial measures as it does not rely on self-ratings, but assesses actions or decisions chosen by the test-taker that are assumed to reveal underlying characteristics.

Examples of talent assessments

ETS's WorkFORCE Assessment for Job Fit

ETS's WorkFORCE Assessment for Job Fit is designed to identify qualified candidates for specific jobs and industries. The assessment measures 6 personality-based behavioural competencies (see Table 19) associated with employee success. It is tailored to the job that an employer is staffing, and an Expected Job Fit Indicator indicates how likely a candidate would succeed at a given position. It is a web-based assessment that takes 20-25 minutes to complete, with 120 pairs of forced-choice items (i.e., test taker must endorse one of two personality statements).

Table 19   **ETS's WorkFORCE Assessment for Job Fit**

Behavioural competencies	Facets	Description
Initiative and perseverance	Diligence, assertiveness, dependability	Individuals higher on this dimension tend to set high standards for themselves and work hard to complete tasks on time and well. They are good at and enjoy leading and influencing others. They tend to be organized with their possessions and tasks.
Responsibility	Dependability, self-discipline, organization	Individuals higher on this dimension typically fulfill their promises and obligations and take responsibility for their mistakes. They are not easily distracted and think things through before acting. They tend to be effective at tasks requiring caution and precision.

Behavioural competencies	Facets	Description
Teamwork and citizenship	Collaboration, generosity	Individuals higher on this dimension tend to trust and get along well with others. They try to be polite, cooperative and **considerate of others' opinions when making decisions.** They tend to be generous with their time, money and possessions. **They are more likely to sympathize with others' problems and** want to help. Such qualities may help these individuals collaborate effectively with colleagues.
Customer service orientation	Collaboration, generosity, friendliness	Individuals higher on this dimension tend to trust and get along well with others. They are more likely to sympathize with **others' problems and want to help. They enjoy talking to** others, feel comfortable meeting new people and consider themselves to have superior social skills. Such qualities may help these individuals effectively interact with customers and **prioritize customers' needs.**
Problem solving and ingenuity	Creativity, intellectual orientation	Individuals higher on this dimension tend to be quick at reading and processing information. They seek out and enjoy opportunities to try or learn about new things, come up with creative ideas and solve mentally challenging problems. They may get bored with easy work assignments that do not challenge them.
Flexibility and resilience	Stability, optimism	Individuals higher on this dimension tend to be more optimistic, have higher self-esteem and react to criticism more positively. They are able to perform well under pressure, laugh off stress and move forward without dwelling on or losing sleep over past mistakes. As a result of these qualities, these individuals may adjust well to changing work environments

Sources: https://www.ets.org/workforce/about/how/ and Naemi, Seybert, Robbins & Kyllonen (2014).

The development of the WorkFORCE assessment is based on the big five model of personality. Twenty-one lower order traits were determined through factor analysis of data from a sample of individuals responding to 7 major personality inventories over 5 years, and analysis of the lexical structure of the inventories (see Drasgow et al., 2012). Personality statements representing these lower order traits were grouped into behavioural composites and mapped onto research-based taxonomies of job performance. The 6 behavioural composites used in the assessment are composites of 12 different facets. A detailed technical reports describes the development and validation of the WorkFORCE assessment and the FACETS™ engine used to deliver it (Naemi, Seybert, Robbins, & Kyllonen, 2014).

ETS's FACETS™ engine has a pool of personality statements and utilizes both forced-choice methodology and computerized adaptive testing to generate test items. The test is adaptive such

that each test item is determined by responses to the previous item. This increases accuracy, reduces the number of items needed, and time to complete the assessment. To reduce socially desirable responding or faking, individuals are presented with two statements that are equal on social desirability and asked to select the one that best applies to themselves.

Examples of forced-choice items

Choose one statement that is MORE LIKE YOU:

        a)   I make decisions only after I have all of the facts

        b)   I try to identify the reasons for my actions

        a)   I look forward to the opportunity to learn and grow

        b)   I don't worry about that that have already happened

Source: https://www.ets.org/s/workforce/videos/29107_workforce.html

Much of the documentation available on validity is focused on the FACETS™ engine and the lower-order trait taxonomy on which the WorkFORCE assessment is based. Multiple studies were conducted with workforce samples, ETS incumbents, and military, in the U.S., and also internationally (PIACC). The FACETS™ engine and trait taxonomy were able to predict a range of academic, work, and organizational outcomes, including turnover, adaptability, contextual performance, task performance, behavioural interview ratings, recruiter ratings of overall fit, advancement in the hiring process, annual performance review ratings, and self-reported SAT scores and GPA.

Pearson's Workplace Personality Inventory

The Workplace Personality Inventory (WPI) is based on the 16 work styles of the O*NET. Work styles describe dispositional or personality work requirements of an occupation, in addition to work requirements related to abilities, occupational interests, and work values. These four elements make up the *worker characteristics* that are important for each occupation. The O*NET work styles taxonomy (Borman, Kubisiak, & Schneider, 1999) was developed by examining existing personality taxonomies (e.g., big five model, Hogan Personality Inventory) and research on personality and job performance.

The 16 work styles used in the WPI are shown in Table 20. It includes some aspects that might overlap with soft essential skills, including critical thinking (i.e., analytic thinking), and working with others (i.e., cooperation, concern for others, social orientation). The WPI is an online/computer assessment with 175 items (8-12 per work style) and takes 20-30 minutes to complete. It was designed to be work-relevant, short, easy-to-understand, and resistant to faking (i.e., items written to minimize social desirability, use of warnings, including "unlikely virtues" scale to identify faking). There is an accompanying guide for employers about how to use WPI for hiring decisions. In addition to profiling the jobs they are filling, they can consult with O*NET job profiles,

which will list relevant work styles, and also use a Work Style Importance Rating Form to
determine the importance of each style to the job.

Table 20 **Pearson's Workplace Personality Inventory work styles**

Work style domain	Work style	Relevant behaviours
Achievement orientation	Achievement/Effort	▪ Establishes challenging goals ▪ Maintains goals ▪ Exerts effort toward task mastery
	Persistence	▪ Persists in the face of obstacles on the job
	Initiative	▪ Takes on job responsibilities without being told to do so ▪ Volunteers for new job responsibilities ▪ Volunteers for new job challenges
Social influence	Leadership Orientation	▪ Demonstrates a willingness to lead/take charge ▪ Demonstrates a willingness to offer opinions
Interpersonal orientation	Cooperation	▪ Is pleasant/good-natured with others on the job ▪ Encourages people to work together ▪ Helps others with tasks
	Concern for Others	▪ Demonstrates sensitivity to the needs and feelings of others ▪ Demonstrates understanding of others/empathy
	Social Orientation	▪ Shows a preference for working with others ▪ Develops personal connections with work colleagues
Adjustment	Self-Control	▪ Keeps emotions in check even in difficult situations
	Stress Tolerance	▪ Accepts criticism ▪ Shows tolerance of stress caused by other people or situations
	Adaptability/Flexibility	▪ Adapts to change in the workplace ▪ Deals effectively with ambiguity ▪ Demonstrates openness to considerable variety in the workplace
Conscientiousness	Dependability	▪ Fulfills obligations reliably
	Attention to Detail	▪ Completes work tasks thoroughly ▪ Is careful about details
	Integrity/Dutifulness	▪ Avoids unethical behaviour ▪ Follows rules and regulations
Independence	Independence	▪ Relies mainly on self to get things done ▪ Develops own way of doing things
Practical Intelligence	Innovation	▪ Generates new ideas to address work issues and problems
	Analytic Thinking	▪ Uses logic to address work-related issues ▪ Produces high quality, useful information

Source: p. 9 in NCS Pearson, Inc. (2007b).

There is documentation of its development and reliability and validity available online (NCS Pearson, Inc., 2007a, 2007b). Reliability and validity were assessed using workforce samples across various occupations and levels. There is relatively good internal reliability for each scale (i.e., alpha all above .76 except one), and the scales are fairly independent of each other. Convergent validity was demonstrated by expected correlations with other personality inventories, such as the Occupational Personality Questionnaire and the Hogan Personality Inventory. Criterion validity was established by relating WPI results to on-the-job performance of incumbents. WPI work styles were also able to differentiate various job groups (e.g., sales representatives scored highest on Social Orientation, compared to others). When comparing "honest" and "applicant" test administrations (i.e., answering honestly, or answering as if in a job selection situation), scores are inflated under applicant conditions. However, the change is smaller than typically reported in other studies with other personality inventories. The *Unlikely virtues Scale* can be used to assess how likely an individual is misrepresenting himself. High scores on this scale would indicate caution in using WPI results.

Given the proprietary nature of the assessment and high-stakes usage, no further information is available online about the actual test items. A related assessment, the Workplace Personality Inventory II, builds on the original WPI, measuring the same 16 work styles, designed for recruiting and coaching early career professionals. Job-related traits are summarized into 6 broad performance areas: achievement, social influence, interpersonal, self-adjustment, conscientiousness, and practical intelligence.

ACT's WorkKeys Talent Assessment

ACT's WorkKeys Talent Assessment is designed for prescreening and selecting employees, and for identifying training needs and developing employees. The assessment can be taken online and usually takes 30 to 35 minutes to complete. There are 165 items that are rated on 6-point rating scale, measuring 12 talent (personality) scales, as well as four talent indices. These are shown in Table 21 and Table 22.

Table 21   **ACT's WorkKeys talent assessment scales**

Big 5 domain	Talent scales	Description	Sample item
Conscientiousness	Carefulness	The tendency to think and plan carefully before acting or speaking	I prefer to plan ahead
	Discipline	The tendency to be responsible, dependable, and follow through with task without becoming distracted or bored	Once I start a task, I see it through to the end
	Order	The tendency to be neat and well-organized	My workspace is usually cluttered (reverse keyed)

Big 5 domain	Talent scales	Description	Sample item
Emotional stability	Stability	The tendency to maintain composure and rationality in situations of actual or perceived stress	I am stressed easily (reverse keyed)
	Optimism	The tendency toward having a positive outlook and confidence in successful outcomes	I tend to believe that things will work out for the best
Agreeableness	Cooperation	The tendency to be likeable and cordial in interpersonal situations	I usually get along well with others
	Goodwill	The tendency to be forgiving and to believe that others are well-intentioned	I think most of the people I deal with are selfish (reverse keyed)
Extraversion	Sociability	**The tendency to enjoy being in other people's** company and to work with others	I frequently attend social gatherings
	Influence	The tendency to impact and dominate social situations by speaking without hesitation and often becoming a group leader	I am often the leader of groups I belong to
	Striving	The tendency to have high aspiration levels and to work hard to achieve goals	I know what my goals are and I constantly work toward them
Openness	Creativity	The tendency to be imaginative and to think **"outside the box"**	I enjoy finding creative solutions to problems
Multiple traits and Emotional Intelligence	Savvy	**The tendency to read other people's motives,** understand office politics, and anticipate the needs and intentions of others	It is easy for me to pick up on the politics at work

Source: p. 5 in ACT, Inc. (2009).

Table 22 **ACT's WorkKeys talent indices**

Talent indices	Description
Teamwork	Refers to the extent to which an individual will demonstrate compromise, cooperation, and interpersonal understanding when working in teams
Work discipline	Refers to the extent to which an individual will demonstrate dependability, as well as disciplined and positive attitude toward the job, rules and regulations, and the work environment
Managerial potential	Refers to the potential that an individual will demonstrate a high level of work performance in supervisory/managerial roles
Customer service orientation	Refers to the potential that an individual will demonstrate a high level of attentiveness, courtesy, and helpfulness in serving customers

Source: p. 7 in ACT, Inc. (2009).

The talent scales were developed using the personality and organizational literature and expert review to identify key constructs related to job performance (ACT, Inc., 2009). These key constructs were used generate an initial pool of items, which were then reduced or revised using readability ratings and assessments by experts and by a sample of employees. A final set of items was determined based on exploratory and confirmatory factor analysis and correlation between employee responses to items and matched supervisor ratings. Validity analyses show good internal consistently (alpha is .81 to .89) and a good fit to the big five model of personality. The talent scales correlated well with supervisor ratings of task performance, prosocial/organizational citizenship, counterproductive behaviours, safety behaviours, communication, teamwork behaviours, productivity, and high performance.

The four talent indices are composites created using items from the 165-item set. A theoretical model for job performance was identified using the literature, comprised of the four constructs of teamwork, work discipline, managerial potential, and customer service orientation. To create each talent index, items from the original 165 item talent scales were selected based on their relevance to each construct. Items were also selected based on how well employee responses to the items correlated with supervisor ratings of the construct. Validity analyses show good internal consistency for the indices (alpha is .85 to .89), and each index correlates well with supervisor ratings. Convergent and discriminant validity was demonstrated by expected patterns of correlations with other job performance criteria (e.g., task performance, counterproductive behaviours) and with the big five domains of personality.

Other work-related personality inventories

There are several other well-known personality inventories that have been developed by individuals or smaller testing companies to be used in organizational and occupational contexts. Several of these are described briefly in Table 23 below. Like the examples above, these inventories are commercial and have been developed based on occupational needs and validated using workforce samples.

Table 23   Examples of other work-related personality inventories

Instrument	Description	Structure
Hogan Personality Inventory (HPI) *Robert Hogan & Joyce Hogan*	Commercial Designed for use in organizations 206 True/False items 15-20 minutes to complete Based on five factor model	7 primary scales, 6 occupational scales, and 42 subscales: Adjustment (empathy, not anxious, no guilt, calmness, even tempered, no complaints, trusting, good attachment) Ambition (competitive, self-confidence, accomplishment, leadership, identity, no social anxiety) Sociability (likes parties, likes crowds, experience seeking, exhibitionistic, entertaining) Interpersonal Sensitivity (easy to live with, sensitive, caring, likes people, no hostility) Prudence (moralistic, mastery, virtuous, not autonomous, not spontaneous, impulse control, avoids trouble) Inquisitive (science ability, curiosity, thrill seeking, intellectual games, generates ideas, culture) Learning Approach (education, math ability, good memory, reading) 6 occupational scales: Service Orientation, Stress Tolerance, Reliability, Clerical Potential, Sales Potential, Managerial Potential
Occupational Personality Questionnaire (OPQ32) *Peter Saville & colleagues*	Commercial Designed for use in organizations for employee selection, development etc. Maximum of 104 sets of items, forced-choice format Takes 25 minutes to complete A series of questionnaires, of which OPQ32 is the latest version	32 personality characteristics: Relationships with People: ▪ Influence: persuasive, controlling, outspoken, independent minded ▪ Sociability: outgoing, affiliative, socially confident ▪ Empathy: modest, democratic, caring Thinking Style: ▪ Analysis: data rational, evaluative, behavioural ▪ Creativity and change: conventional, conceptual, innovative, variety seeking, adaptable ▪ Structure: forward thinking, detail conscious, conscientious, rule following Feelings and Emotions: ▪ Emotions: relaxed, worrying, tough-minded, optimistic, trusting, emotionally controlled ▪ Dynamism: vigorous, competitive, achieving, decisive

Instrument	Description	Structure
Gordon Personal Profile-Inventory (GPP-I) *Leonard V. Gordon*	Commercial Designed to assess candidate fit for job selection, job alignment and performance 38 sets of items, forced-choice method Takes 30-35 minutes	9 personality traits: Profile: ▪ Ascendance ▪ Responsibility ▪ Stress tolerance ▪ Sociability ▪ Self-esteem (sum of other four) Inventory ▪ Cautiousness ▪ Original thinking ▪ Personal relations ▪ Vigor

Conclusion

The above examples of personality inventories, research instruments, work readiness assessments, and talent assessments illustrate the range of approaches that can be taken in developing and selecting psychosocial measures. Some of these measures are based on personality theory and models, some are driven by industry and employer needs, and others attempt to integrate both personality and job performance domains. These examples also illustrate the range of contexts in which psychosocial measures are being used, including assessing and developing work readiness skills in graduating high school students and in adults, evaluating the effects of skills and employment training services, and selecting candidates for specific job profiles. Psychosocial measures are also widely used in organizational research, and psychosocial skills are gaining prominence in our education system as important skills to foster and measure for students entering the future economy and labour market.

The value added from augmenting the ES framework with psychosocial skills comes not only from the increasing use in education, work, and industry, but from a strong evidence base for the relevance of these skills. Employers value these skills, sometimes over cognitive and core essential skills, and there is significant evidence demonstrating its ability to predict academic and work outcomes from research studies and validation studies of personality inventories and talent assessments. Methodologically, existing well validated psychosocial and personality measures can serve as alternate or supplemental assessments of soft essential skills, which are typically difficult to measure objectively. The ES framework already includes some key psychosocial skills and by expanding on these, the full range of cognitive and psychosocial skills that underlie socioeconomic success can be covered.

Motivated by the emerging evidence base on the links between psychosocial skills and life outcomes, large-scale international studies have begun to include measures of personality and other non-cognitive skills to complement literacy and essential skills measures. For example, the Organization for Economic Co-operation and Development (OECD) is currently involved in a number of initiatives that recognize the importance of measuring psychosocial skills, such as: i) a longitudinal study of social and emotional skills; ii) incorporating non-cognitive skills measures in the Programme for International Student Assessment (PISA); and iii) establishing an expert group to identify the non-cognitive skills to be included in the upcoming cycle of the Programme for the International Assessment of Adult Competencies (PIAAC).

Phase 1 summary and recommendations: Essential Skills assessments as part of broader performance frameworks

Assessments for different training contexts and target populations

This report provides an inventory of assessment tools for all nine Essential Skills and divides tools into several categories, e.g., 1) those that assess generic skills that may be applied in a variety of work and non-work contexts versus those that are contextualized within an industry-specific set of job performance requirements; and 2) those based on a bank of test items with objectively correct answers versus those based on self- or other-rated skills. Within each category, assessments vary widely with respect to factors such as test length, cost to purchase and administer, and how rigorously they have been validated and have established psychometric properties.

Selecting tools for assessment depends on how well all these factors align with the context in which training is to be delivered, characteristics of the target population, and high or low stakes implications of test results. In general, Essential Skills training should address a clearly identified set of needs, and contribute to the achievement of goals that are meaningful to all stakeholders (learners, trainers, and employers). Learners and trainers will be motivated to engage with assessment tools to the extent that these tools are clearly connected to learning activities aligned with needs, and that results are viewed as representing meaningful outcomes along a clearly-defined training pathway.

Figure 28 illustrates three broad training contexts for which assessment tools may be selected, with different levels of learner need defined according to distance from the labour market. Learners in the first "pre-employment" block typically have multiple barriers that may prevent them from engaging with employment training. In this context, learning activities may focus on boosting transferable Essential Skills that are important for engagement in subsequent training that may eventually lead to employment. For example, one of the service providers in SRDC's Pay for Success project delivered 'readiness to learn' training to clients with multiple barriers, to allow them to transition between services focused on life stabilization and those that offered job search assistance. In this context, standardized generic assessments of core skills (document use and numeracy) were used as part of a framework that also included self-efficacy and well-being measures to document training gains.

Learners in the second block of Figure 28 are typically unemployed or under-employed, but ready to engage with employment training. In this context, learning activities may focus on identifying and bridging specific employability or skills gaps. For example, SRDC's Foundations project's training model was designed to allow participants with different occupational goals to identify their own Essential Skills in relation to those required by their target occupations, and follow customized learning pathways designed to close gaps between observed and required skills. Though the model was informed by pre-existing occupational skill profiles, because of the wide range of targeted occupations and sectors there was no attempt to refine skill requirements with direct input from employers or industry stakeholders. As a result, standardized generic assessments of core skills

(reading, document use, and numeracy) were used to identify needs ad monitor progress, along with other key indicators such as career pathfinding and job search self-efficacy. In addition, a long-term follow-up survey established to what extent participants were continuing to use core skills in their every day activities.

Figure 28   Training contexts and target populations along the employment pathway
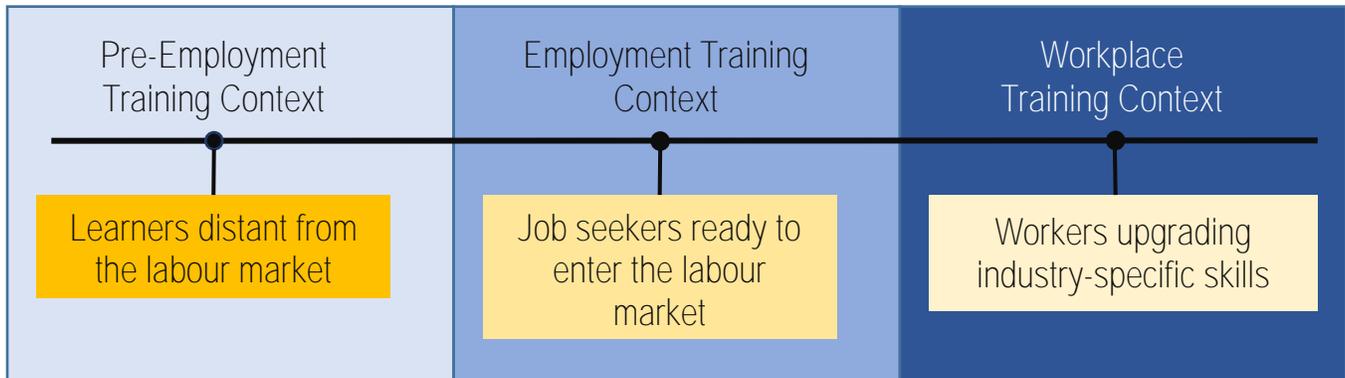
In some cases, learners may transition from one training context to another. For example, one of the Pay for Success providers focused on preparing unemployed job seekers for a sector-based work placement. The first classroom-based training component involved enhancing various employability factors – including core Essential Skills assessed through a standardized generic tool as well as an array of other indicators such as career adaptability and receptivity to continuous learning – with the aim of boosting clients' chances of engaging successfully in the subsequent workplace-based training component. This second component was delivered by sector-based trainers, and involved an industry-contextualized, employer-scored assessment based on learner performance of basic job tasks aligned with eight of the nine Essential Skills.

In general, industry-contextualized assessments are most appropriate for tracking progress in sector-based training models. These kinds of models are typically designed for employed workers seeking to upgrade their job performance, and are delivered in a workplace context (e.g., the third block in Figure 28). In this context, learning activities may focus on embedding Essential Skills training within a well-defined job performance framework. For example, SRDC's New Brunswick cluster training project focused on upgrading the skills of workers in the tourism industry. The model's key objective was to achieve tight alignment between training activities and performance gaps identified by employers and other stakeholders within the tourism sector. There was no appetite for standardized testing – instead, assessment tools were designed, with substantial industry engagement and input, to measure self-reported gains in a range of core and soft Essential Skills underlying improved performance on specific job tasks.

Another workplace example, SRDC's UPSKILL project, took industry-contextualized assessment one step further by embedding Essential Skills training within the full range of business needs of interest to employers in the tourism sector, and measuring progress through a set of industry-wide, expert-assessed standards for certification. In addition, standardized generic assessments were

used to document the gains in core Essential Skills, and self-reported measures were used to capture participants' self-efficacy and frequency of applying a range of Essential Skills both at work and in contexts outside the workplace.

Phase 1 recommendations and next steps

A characteristic common to all the training models described above is that they all used Essential Skills assessments as a part of a broader conceptual and measurement framework that linked training goals and expectations with a variety of short- and long-term outcomes based on learner needs. Among other things, this had the effect of ensuring that assessment results intended to have low-stakes implications (e.g., to track program impact at a group level, and be used as a individual diagnostic tool only in conjunction with other indicators) were perceived as such by stakeholders.

When a single set of assessments is used as the only or main success indicator across a range of different training contexts, without aligning the assessments with learner needs, learning activities or expected outcomes, there is a real risk of trainer and learner disengagement. A pertinent example is the application of standardized assessments of core skills in the context of Ontario's Literacy and Basic Skills (LBS) programs, which use different training models to help learners with a range of needs (including official language minority learners, community and postsecondary learners, Indigenous learners, and participants who are deaf) achieve a variety of goals. Though assessment results showed positive gains at a group level, a substantial proportion of individual learners and trainers reported negative experiences, most of which arose from interpreting the results as having high-stakes implications – for example that a low score meant 'failing the test' and lacking the qualifications required for most jobs. In reality, individual scores were not measured precisely enough to be interpreted in this way. Ultimately, use of the assessment was discontinued (the project is described more fully in Box 1 on p. 56-57).

In general, it is best to incorporate assessments as part of a broader milestone-based performance framework, with multiple interconnected indicators of success. An example of such a framework, from SRDC's Pay for Success project, is shown in Figure 29. This framework defines a sector-based pathway for unemployed job seekers with the aim of transitioning them from classroom-based employability training to work placement and technical training, with the ultimate goal of sustained employment. Expected outcomes are defined at each stage as numbered milestones.

Early success indicators (e.g., milestone 1 and 2 in Figure 29) are defined as gains in core Essential Skills assessed with standardized tools, but also a range of self-reported measures such as career adaptability and receptivity to continuous learning. Results showed that gains in both kinds of measures – standardized assessments and self-reported scales – were linked with successful transitions to workplace training. For those who transitioned to the workplace, performance on an employer-rated assessment of Essential Skills embedded within common job tasks (milestone 3) was linked with the attainment of objectively scored industry-wide productivity standards (milestone 4), and ultimately with employment and job retention (milestones 5 through 8). In addition, the framework includes a variety of contextual indicators to test to what extent the training was suitable for different target populations (e.g., women compared to men, those with higher vs. lower levels of education).

Figure 29   Sample pathway for an industry-specific employment readiness training program
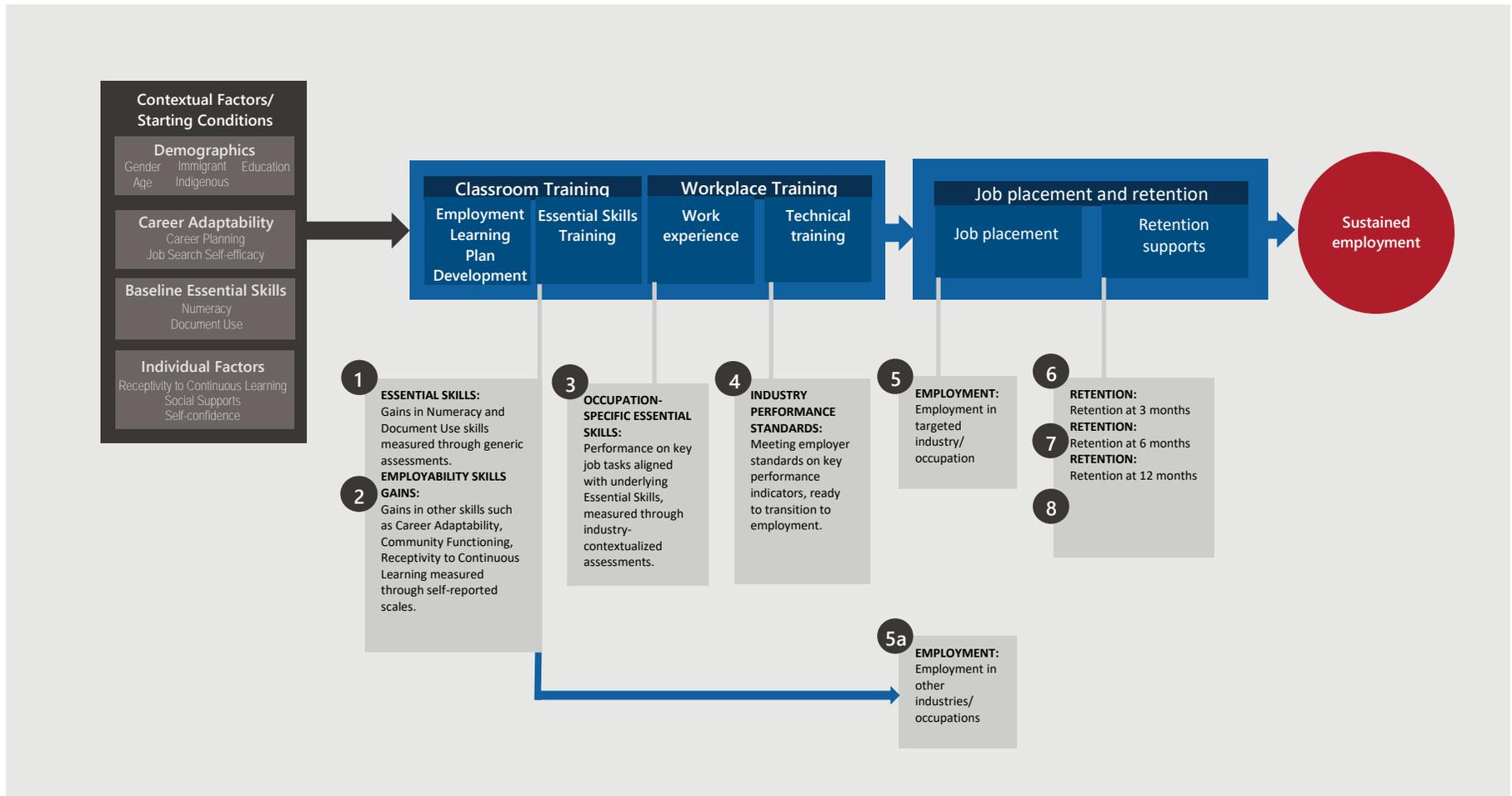
Our goal in Phase 2 is to develop menus of measurement options that integrate:

1. Assessments of core and soft Essential Skills

2. Other important skills and attributes outside the ES framework, such as psychosocial skills

3. Key short-, medium- and long-term indicators of downstream success, such as transition to employment or further education

4. Contextual factors that may help to identify conditions for success, and explain why specific training models may have worked better for some people than others.

We will also develop guidelines to allow various stakeholders to use these menus to build milestone-based performance frameworks contextualized to the needs and training goals of a range of different target populations. Figure 30 offers examples of potential milestones, including in-program gains in Essential Skills and psychosocial skills as well as indicators of post-training success, for each of three broad training contexts. The figure also illustrates several possible contextual factors that may be used to gauge training success for a variety of sub-populations.

SRDC will develop the menus and guidelines for Phase 2 performance frameworks by building on Phase 1 findings, while continuing to draw on evidence from earlier SRDC projects such as Pay for Success, Foundations, and UPSKILL. In addition, we will consult with ESDC to identify other applications of performance frameworks with integrated Essential Skills assessments, and conduct additional key informant interviews as needed.

Figure 30   Sample pathway for an industry-specific employment readiness training program
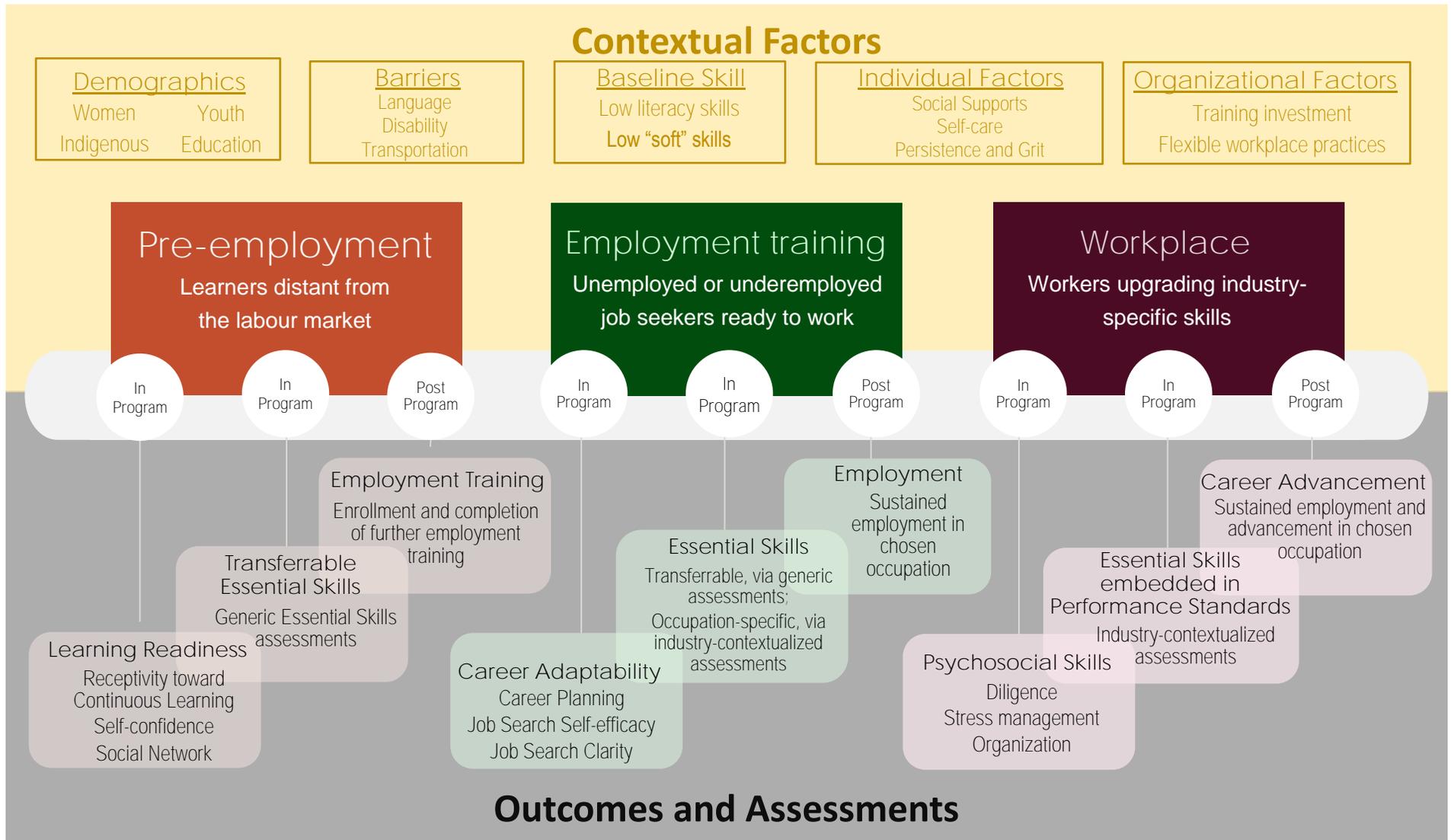
References

ACT, Inc. (2009). *WorkKeys Talent Assessment user and technical guide.* Author.

Almlund, M., Duckworth, A. L., Heckman, J. J., & Kautz, T. (2011). *Personality psychology and economics* (Discussion Paper Series No. 5500). Retrieved from IZA Institute of Labor Economics website: https://www.iza.org/publications/dp/5500/personality-psychology-and-economics

APA. (2018). *APA dictionary of psychology*. Retrieved from https://dictionary.apa.org/

Ashton, M. C. (1998). Personality and job performance: The importance of narrow traits. *Journal of Organizational Behavior, 19*(3), 289-303.

Barrick, M. R., & Mount, M. K. (1991). The Big Five personality dimensions and job performance: A meta-analysis. *Personnel Psychology*, *44*(1), 1–26.

Barton, P. (2006). *High school reform and work: Facing labour market realities*. Princeton, NJ: Educational Testing Service.

Borman, W. C., Kubisiak, U. C., & Schneider, R. J. (1999). Work styles. In N. G. Peterson, M. D. Mumford, W. C. Borman, P. R. Jeanneret, & E. A. Fleishman (Eds.), *An occupational information system for the 21st century: The development of O*NET* (p. 213-226). Washington, DC: American Psychological Association.

Burrus, J., Jackson, T., Xi, N., & Steinberg, J. (2013). *Identifying the most important 21st century workforce competencies: An analysis of the Occupational Information Network (O*NET)* (Research Report ETS RR-13-21). Retrieved ETS website: https://www.ets.org/Media/Research/pdf/RR-13-21.pdf

Canada (1991), Adult literacy in Canada: results of a national study. Labour and Household Surveys Division, Statistics Canada. Catalogue Number 89-525E. Ottawa.

Cartwright, F., TOWES, & Murray, S. (2014) Accuracy and Validity in TOWES Assessments. Unpublished document.

Chen, G., Gully, S. M., & Eden, D. (2001). Validation of a New General Self-Efficacy Scale. *Organizational Research Methods, 4*(1), 62-83.

Clark, B., Hennessy, S., Herzog, M. (2016). Learner Gains Research Project Final Report. Prepared for the Ministry of Advanced Education and Skills Development by the College Sector Committee for Adult Upgrading. Accessed online September 2018: http://www.tcu.gov.on.ca/eng/eopg/publications/lbs-lgrp-report-learner-gains-research-csc-en.pdf

Cronbach, L. J., & Gleser, G. C. (1957). Psychological tests and personnel decisions, 2nd ed. Urbana, IL: University of Illinois.

Costa, P. T., & McCrae, R. R. (1992). Four ways five factors are basic. *Personality and Individual Differences, 13*(6), 653-665.

Duckworth, A., & Quinn, P. D. (2009). Development and validation of the Short Grit Scale (GRIT-S). *Journal of Personality Assessment, 91*(2), 151-174.

Drasgow, F., Stark, S., Chernyshenko, O. S., Nye, C. D., Hulin, C. L., & White, L. A. (2012). *Development of the Tailored Adaptive Personality Assessment System (TAPAS) to support army selection and classification decisions* (Technical Report 1311). Retrieved from http://www.dtic.mil/dtic/tr/fulltext/u2/a564422.pdf

Egan, M., Daly, M., Delaney, L., Boyce, C. J., & Wood, A. M. (2017). Adolescent conscientiousness predicts lower lifetime unemployment. *Journal of Applied Psychology, 102*(4), 700-709.

Ellingson, J. E., Sackett, P. R., & Hough, L. M. (1999). Social desirability corrections in personality measurement: Issues of applicant comparison and construct validity. *Journal of Applied Psychology, 84*(2)*,* 155-166.

Goldberg, L. R. (1971). A historical survey of personality scales and inventories. In P. McReynolds (Ed.), *Advances in psychological assessment* (Vol. 2) (p. 293-336). Palo Alto, CA: Science and Behavior Books.

Goldberg, L. R. (1993). The structure of phenotypic personality traits. *American Psychologist, 48*(1), 26-34.

Hanover Research. (2011). *A crosswalk of 21st century skills.* Retrieved from http://www.montgomeryschoolsmd.org/uploadedFiles/about/strategicplan/21stCenturySkills.pdf

Harvey, R.J. & Hammer, A.L. (1999). Item Response Theory. The Counseling Psychologist. Vol 27, Issue 3, p. 353 – 383. https://doi.org/10.1177/0011000099273004

Heckman, J. J., Humphries, J. E., & Kautz, T. (2014). Who are the GEDs? In J. J. Heckman, J. E. Humphries, & T. Kautz (Eds.), *The myth of achievement tests: The GED and the role of character in American life.* Chicago: University of Chicago Press.

Heckman, J.., Malofeeva, L., Pinto, R., & Savelyev, P. (2010). *Understanding the mechanisms through which an influential early childhood program boosted adult outcomes.* Unpublished manuscript. Department of Economics, University of Chicago.

Hillage, J., Regan, J., Dickson, J., & McLouglin, K. (2002). *Employers skill survey 2002* (Research Report RR372). Department for Education and Skills. Retrieved from http://dera.ioe.ac.uk/4589/1/RR372.pdf

Hogan, J., Barrett, P., & Hogan, R. (2007). Personality measurement, faking, and employment selection. *Journal of Applied Psychology, 92*(5), 1270–1285.

Hogan, J., & Roberts, B. W. (1996). Issues and non-issues in the fidelity-bandwidth trade-off. *Journal of Organizational Behavior, 17*(6), 627–637.

Holzer, H. (1997). Is there a gap between employer skill needs and the skills of the work force? In A. Lesgold, M. Feuer, & A. Black (Eds.), *Transitions in work and learning: Implications for Assessment* (p. 6-33). Washington, DC: National Academy Press.

John, O. P. (1990). The "Big Five" factor taxonomy: Dimensions of personality in the natural language and in questionnaires. In L. A. Pervin & O. P. Johns (Eds.), *Handbook of personality: Theory and research* (p. 60-100). New York, NY: Guilford Press**.**

Judge, T. A., Erez, A., Bono, J. E., & Thoresen, C. J. (2003). The Core Self-Evaluations Scale: Development of a Measure. *Personnel Psychology, 56*(2), 303-331.

Kautz, T., Heckman, J. J., Diris, R., ter Weel, B., & Borghans, L. (2014). *Fostering and measuring skills: Improving cognitive and non-cognitive skills to promote lifetime success.* Retrieved from OECD website: http://www.oecd.org/education/ceri/Fostering-and-Measuring-Skills-Improving-Cognitive-and-Non-Cognitive-Skills-to-Promote-Lifetime-Success.pdf

Kirsch, I. S. (2001), THE INTERNATIONAL ADULT LITERACY SURVEY (IALS): UNDERSTANDING WHAT WAS MEASURED. ETS Research Report Series, 2001: i-61. doi:10.1002/j.2333-8504.2001.tb01867.x

Kirsch, I. S., Jungeblut, A., & Mosenthal, P. B. (1998). The measurement of adult literacy. In T. S. Murray, I. S. Kirsch, & L. Jenkins (Eds.), Adult literacy in OECD countries: Technical report on the first international adult literacy survey. Washington, DC: U.S. Department of Education, National Center for Education Statistics.

Levenson, H. (1973). Multidimensional locus of control in psychiatric patients. *Journal of Consulting and Clinical Psychology, 41*(3), 397-404.

McCrae, R. R., & Costa, P. T., Jr. (1996). Toward a new generation of personality theories: theoretical contexts for the five-factor model. In J. S. Wiggins (Ed.), *The Five-Factor model of personality: Theoretical perspectives* (p. 51-87). New York: Guilford Press.

McFarland, L. A. (2003). Warning against faking on a personality test: Effects on applicant reactions and personality test scores. *International Journal of Selection and Assessment, 11*(4), 65–276.

Meek, M. (1991) On Being Literate, London, Bodley Head.

Meriac, J. P., Woehr, D. J., Gorman, C. A., & Thomas, A. L. E. (2013). Development and validation of a short form for the multidimensional work ethic profile. *Journal of Vocational Behavior, 82*(3), 155-164.

Miller, M. J., Woehr, D. J., & Hudspeth, N. (2001). The meaning and measurement of work ethic: Construction and initial validation of a multidimensional inventory. *Journal of Vocational Behavior, 60*(3), 451-489.

Morgeson, F. P., Campion, M. A., Dipboye, R. L., Hollenbeck, J. R., Murphy, K., & Schmitt, N. (2007). Reconsidering the use of personality tests in personnel contexts. *Personnel Psychology, 60*(3), 683–729.

Naemi, B., Burrus, J., Kyllonen, P. C., & Roberts, R. D. (2012). *Building a case to develop noncognitive assessment products and services targeting workforce readiness at ETS.* Retrieved from ETS website: https://www.ets.org/s/workforce_readiness/pdf/rm_12_23.pdf

Naemi, B., Seybert, J., Robbins, S., & Kyllonen, P. (2014). *Examining the WorkFORCE™ assessment for job fit and core capability of the FACETS™ engine* (Research Report ETS RR-14-32). Retrieved from ETS website: https://www.ets.org/s/research/pdf/RR-14-32.pdf

NCS Pearson, Inc. (2007a). *Workplace Personality Inventory: Evidence of reliability and validity.* Author.

NCS Pearson, Inc. (2007b), *Workplace Personality Inventory: Development.* Author.

OECD (2013) Technical Report of the Survey of Adult Skills (PIAAC).

Ones, D. S., & Viswesvaran, C. (1996). Bandwidth-fidelity dilemma in personality measurement for personnel selection. *Journal of Organizational Behavior, 17*(6), 609-626.

Ones, D. S., Viswesvaran, C., & Schmidt, F. L. (1993). Comprehensive meta-analysis of integrity test validities: Findings and implications for personnel selection and theories of job performance. *Journal of Applied Psychology, 78*(4), 679-703.

Ontario Ministry of Education. (2016). *21st century competencies: Foundation document for discussion. Phase 1: Towards defining 21st century competencies for Ontario.* Retrieved from http://www.edugains.ca/resources21CL/About21stCentury/21CL_21stCenturyCompetencies.pdf

Palczyńska, M. and K. Świst (2016), "Measurement Properties of Non-cognitive scales in the Polish Follow-up Study on PIAAC (POSTPIAAC)", *OECD Education Working Papers*, No. 149, OECD Publishing, Paris. http://dx.doi.org/10.1787/c533e448-en

Paunonen, S. V., Rothstein, M., & Jackson, D. N. (1999). Narrow reasoning about the use of broad personality measures for personnel selection. *Journal of Organizational Behavior, 20*(3), 389–405.

Preparatory Training Programs of Toronto (2007). Communications and Math Employment Readiness Assessment: CAMERA. Test Development Report.

Rammstedt, B., Danner, D., & Lechner, C. (2017). Personality, competencies, and life outcomes: results from the German PIAAC longitudinal study. *Large-scale Assessments in Education 5:2.*

Robins, R. W., Hendin, H. M., & Trzeniewski, K. H. (2001). Measuring global self-esteem: Construct validation of a single-item measure and the Rosenberg Self-Esteem Scale. *Personality and Social Psychology Bulletin, 24*(2), 151-161.

Rosenberg, M. (1965). *Society and the adolescent self-image.* Princeton, NJ: Princeton University Press.

Rotter, J. B. (1966). Generalized expectancies for internal versus external control of reinforcement. *Psychological Monographs: General and Applied, 80*(1), 1-28.

Salgado, J. F., Moscoso, S., Sanchez, J. I., Alonso, P., Choragwicka, B., & Berges, A. (2014). Validity of the five-factor model and their facets: The impact of performance measure and facet

residualization on the bandwidth-fidelity dilemma. *European Journal of Work and Organizational Psychology, 26*(4), 628.

Salgado, J. F., & Táuriz, G. (2014). The Five-Factor Model, forced-choice personality inventories and performance: A comprehensive meta-analysis of academic and occupational validity studies. *European Journal of Work and Organizational Psychology, 23*(1), 3-30.

Schwarzer, R., & Jerusalem, M. (1995). Generalized Self-Efficacy scale. In J. Weinman, S. Wright, & M. Johnston, *Measures in health psychology: A user's portfolio. Causal and control beliefs* (p. 35-37). Windsor, UK: NFER-NELSON.

Schyns, B., & von Collani, G. (2002). A new occupational self-efficacy scale and its relation to personality constructs and organizational variables. *European Journal of Work and Organizational Psychology, 11*(2), 219-241.

Sherer, M., Maddux, J. E., Mercandante, B., Prentice-Dunn, S., Jacobs, B., & Rogers, R. W. (1982). The Self-Efficacy Scale: Construction and validation. *Psychological Reports, 51*(2), 663-671.

Spector, P. E. (1988). Development of the Work Locus of Control Scale. *Journal of Occupational Psychology, 61*(4), 335-340.

Statistics Canada (1996). Reading the future: A portrait of literacy in Canada. Statistics Canada. Catalogue Number 89-551-XPE. Ottawa.

Statistics Canada (1997). Backgrounder on the International Adult Literacy Survey. Surveys and Statistical Programs – Documentation: 89F0094X. Ottawa.

Statistics Canada (2011). The Adult Literacy and Life Skills Survey, 2003 and 2008 Public Use Microdata File: User's Manual. Statistics Canada. Ottawa.

Tangney, J. P., Baumeister, R. F., & Boone, A. L. (2004). High self-control predicts good adjustment, less pathology, better grades, and interpersonal success. *Journal of Personality, 72*(2), 271-324.

Todd. P. E., & Zhang, W. (2018). *A dynamic model of personality, schooling, and occupational choice. Retrieved* from https://cowles.yale.edu/3a/todd-adynamicmodelofpersonalityschoolingandoccupationalchoice-dynamic-model-personality-schooling.pdf

Trapmann, S., Hell, B., Hirn, J.-O. W., & Schuler, H. (2007). Meta-analysis of the relationship between the Big Five and academic success at university. *Zeitschrift fu¨r Psychologie, 215*, 132-151.

Valerie G. Ward Consulting Ltd. & Service-Growth Consultants Inc. (2002, revised 2016). Summary of research on the Employment Readiness Scale. Retrieved from the Employment Readiness Scale website: http://www.employmentreadiness.info/sites/employmentreadiness.info/files/files/Organizations/ERS%20Research%20Results_cdn.pdf

Viswesvaran, C., Deller, J., & Ones, D. S. (2007). Personality measures in personnel selection: Some new contributions. *International Journal of Selection and Assessment, 15*(3), 354–358.

Westwood, A. (2004). Skills that matter and shortages that don't. In C. Warchust, I. Grugulis, & E. Keep (Eds.), *The skills that matter.* New York, NY: Palgrave- Macmillan.

Yamamoto, K. and Kirsch, I. (2002) Scaling TOWES and Linking to IALS. Pulled from http://www.towes.com/media/15736/scaling%20towes%20linking%20to%20ials%20-%20kirsch%20and%20yam.pdf

Zemsky, R. (1997). Skills and the economy: An employer context for understanding the school-to-worktransition. In A. Lesgold, M. Feuer, & A. Black (Eds.), *Transitions in work and learning: Implications for assessment.* Washington, DC: National Academy Press.

Appendix A: Properties of skill assessments

Psychometric scales, including proficiency in a domain, are abstractive and not directly observable. Assessment of one's literacy is basically placing the individual along the continuum of an abstract "literacy proficiency scale" such that any person in the population can be ranked accordingly based on the concept of literacy proficiency. The ordinal attribute of proficiency can only be observed indirectly through literacy tasks. Thus, psychometric assessments are estimations of psychometrics (abstractive) through the observable (existential) outcomes of performing various tasks.

Traditional education testing which uses percentage of correct answers to questions in a test is considered as a psychometric assessment of the competency of the domain. Since there are always some other factors that may affect one's performance in completing assessment tasks, measurement errors are unavoidable. In general, there are two types of measurement error concepts:

- Accuracy/Validity – is the degree of proximity of the measurement results to the true value. Since a psychometric attribute is interpreted and applied under some psychological models of related psychometric attributes, accuracy of an assessment is often referred as validity (and sometimes as constant error) of the assessment. That is, the degree to which an assessment measures what it is supposed to measure.

- Reliability/Precision – is the degree to which repeated measurements under unchanged conditions display the same results. Since measurement errors of repeated measurements are considered as random, it affects the reliability of assessment results and degree of precision of an assessment/test is often termed as the reliability of the assessment/test.

Since measurement errors on are unavoidable, decisions made based on assessment result will be subject to decision errors. As a result, it is important to understand the potential harms caused by decision errors and establish a tolerable level of errors.

Another consideration of adopting any assessment into decision making process is the concern about the usage's fairness. Assessments and tests do not always produce the same accuracy and precision to all test takers. Accuracy and precision of an assessment can be substantially lower among particular groups of the population. Some would consider it unfair if the level of decision errors is higher among certain sub-populations.

Validity

Validity is a measure of how much an assessment tool's results actually reflect the characteristics it is intended to assess, instead of other, possibly related characteristics. Validity can be established by analyzing test content and examining the relationship between test items and external variables.

Test content – An assessment tool's content must comprehensively reflect the constructs it is claiming to test. For example, an assessment examining only Numeracy and Document Use might be a valid measure of those two skills, but would not be a valid measure of Oral Communication. Similarly, a Numeracy assessment could not be considered a true assessment of Numeracy if it

excluded a subdomain (e.g., geometry) which is a key to the definition of Numeracy as an Essential Skill. The content of assessment items in IALS, ALL, PIAAC, TOWES, and ESG on prose, document, numeracy have all been examined for their content validity.

Relationship with external variables – Validity can also be determined through the relationships between test items and other external variables. These relationships can be examined in three ways:

1. The correlation between assessment results and other variables which are known to measure Essential Skills should be positive and substantial. For example, the results of a valid Essential Skills assessment should correlate with other, previously validated assessments.

2. The correlation between assessment results and other variables which are logically unrelated to Essential Skills should be low. For example, the results of an Essential Skills assessment should not be significantly correlated with measures of a person's personality traits: if they were, it might suggest that the test is not truly measuring Essential Skills.

3. The correlation between assessment results and the expected results of Essential Skills should be positive and substantial. For example, if higher literacy or Essential Skills levels are known to predict higher income, then assessment results should be positively correlated with participant income levels.

Precision

An assessment's precision generally refers to the random measurement error associated with its results, or how much assessment results tend to vary around a person's true level of Essential Skills. In its most basic form, precision can be thought of as how much the results of an assessment are susceptible to change if a participant makes a random error or a lucky guess on an item which does not reflect their actual skills. This kind of error could be caused by variety of things, including lack of motivation or interest, momentary inattention, inconsistent application of skills, variation in testing conditions, and variations in scoring. The precision of an assessment affects the statistical power with which associations between individual gains and subsequent outcomes can be identified, as well as the sample size required to accurately identify gains at a group level.

Under the Classical Test Theory framework, precision is usually represented as a reliability constant for the whole assessment, regardless of the individual. Under the Item Response Theory framework, precision is usually calculated in terms of the standard error associated with each participant's score, based on the number of items the participant answered consistently correct at one level of difficulty but incorrect responses at the next level. Standard error can be used to calculate confidence interval associated with the participant's score, or the range within which we can be 95% sure the participant's score lies. Assessments with high individual standard errors cannot be relied upon at an individual level, since the participant's true score could lie within a very wide range.

Unlike validity and fairness, the required level of precision can vary depending on the intended use of the assessment. In general, it is more important for an assessment tool to be precise when the

results are being used at an individual level than when results only being used at population level. We discuss how requirements for precision vary under three different assessment purposes.

1. **Individual diagnostic –** In the case of diagnostic assessment, individual participants are assessed in order to determine their levels of Essential Skills, often to inform their training options. If an imprecise assessment is used, variance from the participant's true level of Essential Skills could cause false positive or false negative results. These could take the form of a participant with a high level of Essential Skills being assessed at a lower level, resulting in unnecessary training, or a participant with a lower level of Essential Skills being assessed at a higher level, resulting in the participant not receiving potentially valuable training. Since the stakes of the assessment outcome may be high, the assessment tool used should have the maximum plausible level of precision.

2. **Program evaluation –** In a program evaluation context, assessment results are not used to identify the Essential Skills level of any one individual, but instead to understand the Essential Skills levels and gains of an entire group participating in a training intervention. Since data is aggregated in this case over all participants in the program, lower precision is acceptable as individual misclassifications will not affect estimates of program impacts when they are averaged across the entire group. Furthermore, the stakes of individual assessments are lower, as they are not being used to make training decisions at an individual level. However, the precision of individual assessments does affect the statistical power to identify the impacts of a program, especially for programs with a smaller number of participants. As a result, assessments of moderate to high precision should be used in program evaluation contexts, depending upon the number of participants in the program and the stakes associated with the evaluation (e.g., if it will be used to determine future funding or programming decisions).

3. **Population-level research** – In research projects profiling the overall Essential Skills of a population or linking the results of an Essential Skills intervention with other outcomes, assessment results are aggregated over a very large sample of the population. This scenario requires the lowest assessment precision of the three, as individual misclassifications are balanced out when results are aggregated. However, at least moderately precise assessments are desirable for more accurate population profiles or to verify the gains produced by an intervention, particularly when sample sizes are smaller. Higher precision would be required for linking the magnitude of Essential Skills gains with other outcomes at an individual level.

Trade-off between precision and assessment length – Precision is always a desirable quality of an assessment tool. However, the most reliable way to increase the precision of an assessment is often to increase its length: longer and more involved assessments provide more opportunity for the true Essential Skills of the participant to be reflected, and minimize the potential variation create by random errors. This results in a trade-off between precision and ease of implementation when selecting assessment tools, as tools which are more precise will often require more time from participants and facilitating staff. In cases of individual diagnostic assessment, longer assessments are often preferable, as the potential costs of misclassifying a participant outweigh the added length of the assessment. However, in cases of program evaluation or population-level research, using

shorter assessments may be a worthwhile way to ensure that the assessment is feasible to implement.

Fairness

Fairness refers to the degree to which the administration and results of an assessment do not favour certain individuals over others due to factors beyond the individual's Essential Skills level. Fairness is important in the context of the proposed study, as any systematic differences in the fairness of the assessment between different population subgroups could lead to incorrect estimates of the effects of Essential Skills upgrading. A fair Essential Skills assessment should fit several criteria:

1. **Fair administration –** The assessment should be administered to all participants in an equitable manner, allowing for participants of a range of skill levels to comfortably complete the assessment.

2. **Equitable access –** The assessment should be fully accessible for all eligible participants. Specifically, it should accommodate participants with disabilities which could affect their ability to complete some assessment components. In addition, the assessment should require clear language and literacy benchmarks which can be used to screen out participants who have not attained the minimum language requirements required to use the assessment.

3. **Unbiased content –** The content of the assessment should not unfairly favour participants of specific gender, age, religion or cultural background.

Fairness is not as easily assessed as validity and precision, as it does not correspond to specific test metrics. However, it is possible to conduct small scale pilots and examine whether certain disadvantaged sub-populations perform poorly in a test due to reasons beyond their proficiencies in literacy and Essential Skills.

Item Response Theory

Competency assessments would have been relatively simple if the tasks of a domain of skills are well defined, identical in real life, and inexpensive to conduct. For example, the competency of an individual in operating a passenger car safely on the public roads can be assessed through a road test. The tasks, situations and the sequences of actions and decisions to handle each task and situation in driving are well-defined and driving for a road test is not different from driving in everyday life. However, this is not the case for assessment of proficiency of some basic skills. For example, one may be able to conduct divisions of numbers on paper (such as by memorization), but the person may not be able to apply divisions to calculate unit pricing while shopping in a shopper market, or to apply divisions in distributing items to a group of people. Assessments of such skills are essentially measuring some latent personality traits (not directly observable) through the responses to assessment items and tasks.

Traditional assessments account for the extent of correct responses (such as calculating the test score as the total number of correct answers in a multiple choice test) as a measure of the competency or proficiency of the person. Under the Classical Test Theory (CTT), a personality trait

is represented by the true score of a particular test/assessment, and the observed score is a measure of the true score with measurement errors. The observed test scores are comparable at different time or for different people as long as the construct is invariant. Although it is possible to make different test or assessment instruments comparable under CTT, one must first establish the linkage properties of different tests through examination of patterns of items to some common anchor test items or tasks. However, without a strong model, the relationships between items to the true score are also sample dependent. The test dependency and sample dependency of classical tests thus make it difficult to conduct norm-referenced comparisons.

It is beyond the scope of this study to provide a detailed comparison of scale construction techniques under CTT and IRT, and interested readers should refer to psychometric literature elsewhere (such as Harvey and Hammer, 1999). The following introduces some basic concepts and properties of IRT to help understanding the construction of a few core essential skills assessments being used in the field.

IRT specifies the relationship between the probability of a person's response to an item conditional on the person's ability. For example, a commonly used IRT model is 2-Parameter Logistic Model (2PL) of ability test. The probability of person *i* of providing a correct answer to item *j* is expressed as

(1) $P(x_j = 1|\theta_i) = 1/[1 + \exp(-a_j(\theta_i - b_j))],$

where θ_i is person *i*'s ability, b_j is the difficulty parameter of item j, a_j is the discrimination parameter that describes how well item j separates people of lower and higher ability. A person with an ability above the difficulty of the item would answer the item correctly with over 50% probability. The item response function (1) could be represented as an item characteristic curve (ICC) graphically. 0 presents some examples of ICCs with three different difficulties and two different discrimination parameters. An easier item (e.g., b=-0.8) is represented by an ICC to the left (compared to b=0) while a more difficult item (b=+0.8) is represented by an ICC to the right. An item with a higher discriminatory power (a=2) would have a steeper ICCs compared to an item with a lower discrimination parameter (a=1).

Examples of Item Characteristic Curves of a 2PL Model

Item Characteristic Curves Examples
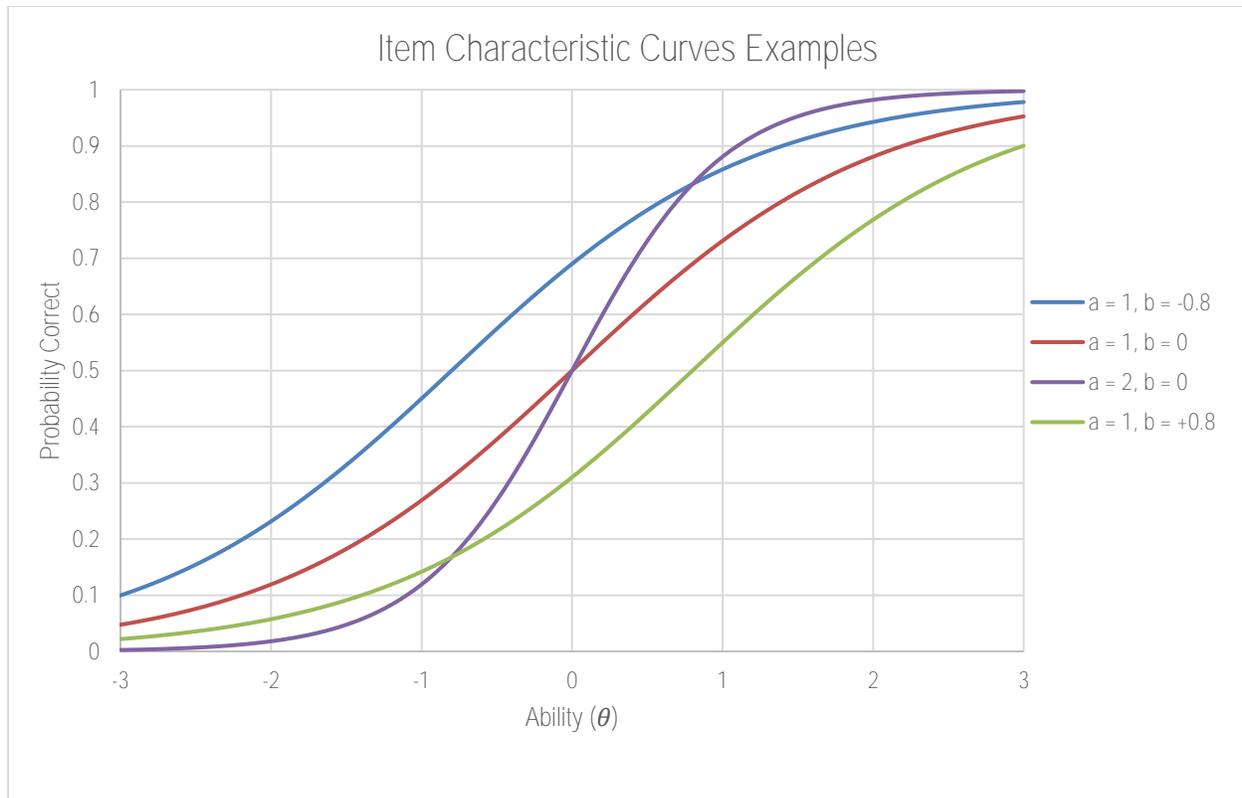
Importantly, the item response functions are used to establish the likelihood function such that item parameters and individual's latent ability are estimated through maximum likelihood estimation from a sample. Usually the distribution of individual ability is assumed to be standard normal though it can be transformed into any range.

With the item response functions, there are many good properties of the model to facilitate revisions of a psychometric assessment instrument:

▪ Each individual's ability estimate is associated with a precision parameter called standard error of measurement. This is in contrast to the reliability parameter (usually the Cronbach's alpha) of a scale constructed under CTT, which is constant for the whole scale. For a fixed set of items, it is common to have a higher standard error of measurement at the low end and high end of ability.

▪ With a sufficiently large number of items, eliminating some items from the scale is likely to affect only the precision of the ability estimate as long as items of similar difficulties remain in the assessment.

▪ The item response function of an item suggests that an item provides the maximum information to an assessment if the item's difficulty matches the person's ability. It implies that an item far away from a person's ability does not have any substantial effect on the accuracy or precision of estimate. This property forms the basis of adaptive/self-level testing. For example, it is possible

to administer the assessment in two steps. The first steps include items covering a range of difficulties in the middle. Respondents who fail to answer all items of the first step correctly will be asked for answers to the least difficult items in the second step while respondents who answer all items of the first step correctly will be asked for the most difficult items. If the question bank is sufficiently large, a computerized adaptive testing can be used to pick the next question maximizing the information based on all collected information until the precision of the assessment meets a pre-determined level.

▪ The sample and test independence of IRT assessment facilitate comparison across different populations and different time without strict requirement of maintaining the same questions.

▪ In practice, it is common that the IRT score is highly correlated with the raw score used in CTT. Indeed, for the one parameter logistic model, IRT score is a monotonic transformation of the raw CTT score. It is therefore possible to cross-reference IRT score with raw CTT score.

The 2PL model can also be extended to handle ordered polytomous data (to account for partial credits in open-ended questions) through a graded response model (GR). The item response function for the GR model is:

(2) $P\big(x_j \geq c | \theta_i\big) = 1/[1 + \exp\big(-a_j(\theta_i - b_j^c)\big)],$

where c is a rating category while b_j^c is the item's category boundary.

These advantageous properties of IRT come with the strong assumptions of unidimensional ability and local independence. The simplest IRT models are not applicable if there is more than one dimension of ability informing the item response. However, recent development has extended IRT models to handle multiple dimensions. IRT estimates would be biased if the responses to any two items were dependent. As a result, construction of IRT assessment requires careful examination of question properties to ensure the data collected fit well with the IRT model.

Task characteristics from NALS and IALS

	Prose	Document	Quantitative
Level 1 (0-225)	Most of the tasks in this level require the reader 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 reader 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 readers to perform single, relatively simple arithmetic operations, such as addition. The numbers to be used are provided and the arithmetic operation to be performed is specified.

	Prose	Document	Quantitative
Level 2 (226-275)	Some tasks in this level require readers to locate a single piece of information in the text; however, several distractors or plausible but incorrect pieces of information may be present or low-level inferences may be required. Other tasks require the reader 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.	Tasks in this level are more varied than those in level 1. Some require the readers to match a single piece of information; however, several distractors may be present or the match may require low-level inferences. Tasks in this level may also ask the reader to cycle through information in a document or to integrate information from various parts of a document.	Tasks in this level typically require readers to perform a single operation using numbers that are either stated in the task or easily located in the material. The operation to be performed may be stated in the question or easily determined from the format of the material (for example, an order form)
Level 3 (276-325)	Tasks in this level tend to require readers to make literal or synonymous matches between the text and information given in the tasks, or to make matches that require low-level inferences. Other tasks ask readers to integrate information from dense or lengthy text that contains no organizational aids such as headings. Readers 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 reader to integrate multiple pieces of information from one or more documents. Others ask readers to cycle through rather complex tables or graphs which contain information that is irrelevant or inappropriate to the task.	In tasks in this level, two or more numbers are typically needed to solve the problem, and these must be found in the material. The operation(s) needed can be determined from the arithmetic relation terms used in the question or directive.
Level 4 (326-375)	These tasks require readers 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 reader.	Tasks in this level, like those at the previous levels, ask readers to perform multiple-feature matches, cycle through documents, and integrate information; however, they require a greater degree of inferencing. Many of these tasks require readers 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 reader.	These tasks tend to require readers to perform two or more sequential operations or a single operation in which the quantities are found in different types of displays, or the operations must be inferred from semantic information given or drawn from prior knowledge.

	Prose	Document	Quantitative
Level 5 (376-500)	Some tasks in this level require the reader to search for information in dense text which contains a number of plausible distractors. Others ask readers to make high-level inferences or use specialized background knowledge. Some tasks ask readers to contrast complex information.	Tasks in this level require the reader to search through complex displays that contain multiple distractors, to make high-level text-based inferences, and to use specialized knowledge.	These tasks require readers to perform multiple operations sequentially. They must disembed the features of the problem from text or rely on background knowledge to determine the quantities or operations needed.

Source: Figure 7.3 (p. 133) in Kirsch, Jungeblut, and Mosenthal (1998).

Appendix B: Essential Skills Indicator – sample items

Numeracy

6. Airline sales agents keep customers informed of changes to flight schedules.

Flight DG386 was originally scheduled to arrive at 13:00. The agent is advised that the flight will be delayed by 2 hours.

What is the new arrival time for Flight DG386?

[   ]

Document Use

2. Window installers complete time sheets.

Look at the time sheet below.

Window Clear Inc. Time Sheet				
Employee:	**Ajit Neelamkavil**			
Manager:	Erik Meinert			
Week Ending:	March 12			
Day	**Date**	**Regular Hours**	**Overtime Hours**	**Total**
Monday	March 6	8	0	8
Tuesday	March 7	8	2	10
Wednesday	March 8	8	0	8
Thursday	March 9	8	3	11
Friday	March 10	8	1	9
Saturday	March 11	4	4	4
Sunday	March 12	0	0	0

How many overtime hours did Ajit Neelamkavil work on Thursday?

[   ]

Reading

2. Nurse aides read emergency procedures.

A patient at a hospital is angry and has thrown a dangerous chemical product onto the floor.

Look at the Emergency Code Manual below.

Emergency Code Manual		
Emergency Code	**Description**	**Immediate Action(s)**
Code Pink	A child has been abducted.	Contact Security to provide a description of the abducted child.
Code White	Hazardous materials have been spilled.	Cordon off the area. Contact maintenance to clean the spill.
Code Yellow	The hospital will be receiving a large number of patients at once.	Activate the disaster plan.
Code Grey	A patient or other individual is being combative.	Contact Security. If possible, contain the individual in an isolated area.
Code Purple	An adult patient is missing.	Contact Security to provide a description of the missing patient.
Code Red	There is a fire in the building.	Pull the fire alarm. Close all doors and windows.

Which two codes should the nurse aid act upon?
- ☐ A. Code Pink
- ☐ B. Code White
- ☐ C. Code Yellow
- ☐ D. Code Grey
- ☐ E. Code Purple
- ☐ F. Code Red

Appendix C: UPSKILL – Self-reported skills

Workplace Context

Confidence in skill application

1. **Please tell me whether you strongly disagree, disagree, agree, or strongly agree with the following statements.**

    Please check off the appropriate box for each statement below.

	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
I have the reading skills in English I need to do my main job well.	O	O	O	O	O
I have the writing skills in English I need to do my main job well.	O	O	O	O	O
I have the oral communication skills (i.e., ability to speak to a supervisor and co-workers) in English that I need to do my main job well.	O	O	O	O	O
I have the math skills I need to do my main job well.	O	O	O	O	O
I have the technical skills I need to do my main job well.	O	O	O	O	O

Frequency in skill usage

2. **The following questions ask about your skills you use specifically <u>in your main job</u>. How often do you <u>read or use information</u> from each of the following as part of your main job?**

    Please check off the appropriate box for each item below.

	Never	Rarely	Less than once a week	Once a week	A few times a week	Every day
Letters, memos or e-mails	O	O	O	O	O	O
Directions, instructions, manuals or reference books (including catalogues)	O	O	O	O	O	O
Reports, bills, invoices, spreadsheets or budget tables	O	O	O	O	O	O

3. How often do you <u>write or fill out</u> each of the following as part of your main job?

Please check off the appropriate box for each item below.

	Never	Rarely	Less than once a week	Once a week	A few times a week	Every day
Letters, memos or e-mails	○	○	○	○	○	○
Directions, instructions, manuals or reference books (including catalogues)	○	○	○	○	○	○
Reports, bills, invoices, spreadsheets or budget tables	○	○	○	○	○	○

4. How often do you do each of the following as part of your main job?

Please check off the appropriate box for each item below.

	Never	Rarely	Less than once a week	Once a week	A few times a week	Every day
Measure or estimate the size or weight of objects	○	○	○	○	○	○
Calculate prices, costs, or budgets	○	○	○	○	○	○
Count or read numbers to keep track of things	○	○	○	○	○	○
Manage time or prepare timetables	○	○	○	○	○	○

Non-workplace context

Confidence in skill application

5. Please indicate the extent to which you agree/disagree with the following statements.

Please check off the appropriate box for each statement below.

	Strongly disagree	Disagree	Neutral	Agree	Strongly agree
I am good with numbers and calculations.	O	O	O	O	O
I feel anxious when figuring out such amounts as discounts, sales tax or tips.	O	O	O	O	O
I read only when I have to.	O	O	O	O	O
Reading is one of my favourite activities.	O	O	O	O	O
I enjoy talking about what I have read with other people.	O	O	O	O	O
I am confident in my ability to write notes, letters or e-mails.	O	O	O	O	O

Frequency of skill usage

6. How often do you do each of the following activities <u>outside of work</u>?

Please check off the appropriate box for <u>each</u> activity below.

	Never	Rarely	Less than once a week	Once a week	A few times a week	Every day
Do math (such as for bills, ban accounts or credit cards)	O	O	O	O	O	O
Read or use information from newspapers or magazines	O	O	O	O	O	O
Read or use information from books – fiction or non-fiction	O	O	O	O	O	O
Read or use information from letters, notes, or e-mails	O	O	O	O	O	O

	Never	Rarely	Less than once a week	Once a week	A few times a week	Every day
Write notes, letters or e-mails	O	O	O	O	O	O
Use a computer outside of work	O	O	O	O	O	O
Use a library or visit a bookstore	O	O	O	O	O	O

Appendix D: OLES Self-assessments for job seekers and workers

Oral Communication

Instructions:

1. Read each statement in **Section 1** and place a check mark in the column that **best** describes how well you can complete that task.

   Tip: Think about your work and life experiences as you consider each task.

2. Review your responses for each task. If you have checked five or more in the "Somewhat" and/or "No" columns, you may want to consider upgrading your **oral communication** skills.

3. Complete **Section 2** to identify your training needs.

Section 1: Self-Assessment

This table contains statements for the oral communication self-assessment.

I can...	Yes	Somewhat	No
Ask routine questions to obtain information.			
Leave brief phone messages.			
Understand short messages and communicate the information to others.			
Give simple instructions to others on a familiar topic.			
Explain simple facts.			
Follow simple oral instructions.			
Listen to others without interrupting.			
Use appropriate body language (e.g., smiling, nodding, making eye contact) while having a conversation.			
Discuss work-related problems or issues in detail.			
Ask complex questions to get the appropriate information.			
Communicate with others to resolve minor conflicts, such as customer complaints.			
Communicate with others to co-ordinate work or resolve problems.			

I can...	Yes	Somewhat	No
Express my opinions and ideas clearly and concisely.			
Restate information that is presented orally.			
Train or give clear instructions to a co-worker.			
Give a brief presentation to a small group.			
Lead routine meetings (e.g., weekly team meetings).			
Follow complex oral instructions to complete a task.			
Explain difficult subject matter using detailed examples.			
Give constructive feedback or advice.			
Speak respectfully to clients or co-workers when dealing with complex issues or resolving conflicts.			
Exchange ideas and opinions with clients, such as clarifying detailed work specifications, or negotiating contracts.			
Persuade others to consider different options.			
Give presentations to a large, unfamiliar group.			
Total			

Section 2: Personal development

Completing this section will help you make informed training decisions.

▪ Look at the "Yes" column in Section 1 to identify your strengths, and record them below.

▪ Look at the "Somewhat" and/or "No" columns in Section 1 to identify the areas that you need to develop or strengthen, and record them below.

Oral communication strengths:

I am confident that I can…

E.g., listen to others without interrupting.

1. _____

2. _____

3. _____

Tip: Consider using your strengths to help a colleague, friend or family member improve their oral communication skills.

Areas for improvement:

I would like to improve my ability to…

E.g., give presentations to a large, unfamiliar group.

1. _____

2. _____

3. _____

Tip: When developing your training plan, focus on improving one or two abilities at a time.

Computer Use

Instructions:

1. Read each statement in **Section 1** and place a check mark in the column that **best** describes how well you can complete that task.

    Tip: Think about your work and life experiences as you consider each task.

2. Review your responses for each task. If you have checked five or more in the "Somewhat" and/or "No" columns, you may want to consider upgrading your **computer use** skills.

3. Complete **Section 2** to identify your training needs.

Section 1: Self-Assessment

This table contains statements for the computer use self-assessment.

I can…	Yes	Somewhat	No
Use a variety of electronic devices such as calculators, fax machines, photocopiers, and telephones.			
Identify the main parts of a computer and their functions.			
Enter a code into a system (e.g., using an automated bank machine or an alarm system).			
Use a mouse to open and navigate programs by clicking buttons, menus, etc.			
Easily locate and use keyboard keys.			
Log on and log off a computer.			

I can...	Yes	Somewhat	No
Use common desktop icons and menus to open programs and files (e.g., Internet browser, email).			
Send a basic email to one recipient.			
Identify and use different computer hardware components such as printers, scanners, and removable storage devices (e.g., CD, DVD, USB drive).			
Resolve basic technical difficulties (e.g., computer reboot, paper jam, ink cartridge replacement).			
Use word processing programs to produce simple documents and perform simple formatting of text (e.g., business letters or memos; font type and size, bold and underline text, bullets or numbered lists).			
Send emails with attachments to multiple recipients.			
Use spreadsheet software to enter and organize data.			
Manage electronic files (e.g., access shared files, save files to a specific location on a hard drive or disk, create folders).			
Browse Web pages to find information and download files.			
Use Internet search engines to find specific information by choosing the right search terms or adding operators (e.g., **OR, " ", define:) to my search terms.**			
Use spreadsheet software to prepare, edit, manipulate and analyze tables (e.g., create and modify budget reports, create various types of charts or graphs).			
Use word processing programs to produce documents with extensive formatting features (e.g., page numbering, headers and footers, table of contents, footnotes).			
Use presentation software to produce visual aids (e.g., slides, hand-outs, and/or charts).			
Give detailed computer-related explanations or training to co-workers (e.g., demonstrating features of commonly used or customized programs).			
Distinguish between different file formats (e.g., rtf, pdf, html, exe, jpg, mpeg).			
Successfully install and configure new software applications or add hardware devices.			
Identify and correct hardware or software problems by using manuals, on-line resources, etc.			
Use and create databases to manage client contact information.			
Total			

Section 2: Personal development

Completing this section will help you make informed training decisions.

▪ Look at the "Yes" column in Section 1 to identify your strengths, and record them below.

▪ Look at the "Somewhat" and/or "No" columns in Section 1 to identify the areas that you need to develop or strengthen, and record them below.

Computer use strengths:

I am confident that I can…

E.g., use common desktop icons and menus to open programs and files.

1. _____

2. _____

3. _____

Tip: Consider using your strengths to help a colleague, friend or family member improve their computer use skills.

Areas for improvement:

I would like to improve my ability to…

E.g., install and configure new software applications or add hardware devices.

1. _____

2. _____

3. _____

Tip: When developing your training plan, focus on improving one or two abilities at a time.

Writing

Instructions:

1. Read each statement in **Section 1** and place a check mark in the column that **best** describes how well you can complete that task.

   Tip: Think about your work and life experiences as you consider each task.

2. Review your responses for each task. If you have checked five or more in the "Somewhat" and/or "No" columns, you may want to consider upgrading your **writing** skills.

3. Complete **Section 2** to identify your training needs.

Section 1: Self-Assessment

This table contains statements for the writing self-assessment.

I can...	Yes	Somewhat	No
Write short reminder notes to myself or others.			
Write simple lists (e.g., grocery list, list of work tasks).			
Write notes in point form.			
Write entries in appointment calendars.			
Write brief notes or emails to co-workers.			
Write brief entries into forms (e.g., bank deposit slip).			
Write down telephone messages.			
Write short memos or faxes to request or clarify information.			
Write text that is a paragraph or longer such as memos or letters.			
Prepare written materials using templates.			
Use correct grammar and spelling.			
Include details and examples to support my writing.			
Tailor my writing to a specific audience.			
Write formal emails or letters to clients or supervisors.			
Write a brief summary of a larger piece of text.			
Write short reports about a meeting or presentation.			
Write long pieces of text such as letters or reports.			
Thoroughly express my opinions in writing.			
Prepare comprehensive written materials to provide direction, training, or support (e.g., a training manual).			
Write detailed emails or letters requesting information.			
Write long pieces of text using formatting features such as headings, table of contents, footnotes, etc.			

I can...	Yes	Somewhat	No
Prepare complex written materials (e.g., contracts, financial reports) using a standard format.			
Rewrite text to improve clarity and flow.			
Edit written materials to confirm proper grammar, spelling, and formatting.			
Total			

Section 2: Personal development

Completing this section will help you make informed training decisions.

- Look at the "Yes" column in Section 1 to identify your strengths, and record them below.
- Look at the "Somewhat" and/or "No" columns in Section 1 to identify the areas that you need to develop or strengthen, and record them below.

Writing strengths:

I am confident that I can…

E.g., write down telephone messages.

1. _____

2. _____

3. _____

Tip: Consider using your strengths to help a colleague, friend or family member improve their writing skills.

Areas for improvement:

I would like to improve my ability to…

E.g., tailor my writing to a specific audience.

1. _____

2. _____

3. _____

Tip: When developing your training plan, focus on improving one or two abilities at a time.

Reading

Instructions:

1. Read each statement in **Section 1** and place a check mark in the column that **best** describes how well you can complete that task.

   Tip: Think about your work and life experiences as you consider each task.

2. Review your responses for each task. If you have checked five or more in the "Somewhat" and/or "No" columns, you may want to consider upgrading your **reading** skills.

3. Complete **Section 2** to identify your training needs.

Section 1: Self-Assessment

This table contains statements for the reading self-assessment.

I can...	Yes	Somewhat	No
Read and understand familiar names, words and simple sentences.			
Scan a short document, such as an email, memo or bulletin, and understand its meaning.			
Read and follow directions on a product label (e.g., on a prescription bottle).			
Read a paragraph to find a key piece of information.			
Read a catalogue to learn basic product information.			
Read and correctly follow written instructions (such as a recipe or job assignment).			
Read a product warning label and understand how to handle the product safely.			
Read several short documents, such as brochures or newspaper articles, to find new information (e.g., reading about a travel destination before a trip).			
Read and understand formal documents, such as a credit card agreement, employment contract or insurance policy.			
Refer to several resources, such as handbooks and manuals, to use a new piece of equipment (e.g., fax machine, printer, cellular telephone, dishwasher).			
Read a newspaper editorial and understand the issue.			
Refer to a variety of documents to compare information (e.g., product specifications, features and costs of different models of digital cameras).			
Refer to appropriate resources, such as policy or procedural manuals, when dealing with unfamiliar or unusual problems at work.			
Total			

Section 2: Personal development

Completing this section will help you make informed training decisions.

- Look at the "Yes" column in Section 1 to identify your strengths, and record them below.
- Look at the "Somewhat" and/or "No" columns in Section 1 to identify the areas that you need to develop or strengthen, and record them below.

Reading strengths:

I am confident that I can…

E.g., read and follow directions on a product label.

1. _____
2. _____
3. _____

Tip: Consider using your strengths to help a colleague, friend or family member improve their reading skills.

Areas for improvement:

I would like to improve my ability to…

E.g., refer to a variety of documents to compare information.

1. _____
2. _____
3. _____

Tip: When developing your training plan, focus on improving one or two abilities at a time.

Document Use

Instructions:

1. Read each statement in **Section 1** and place a check mark in the column that **best** describes how well you can complete that task.

   Tip: Think about your work and life experiences as you consider each task.

2. Review your responses for each task. If you have checked five or more in the "Somewhat" and/or "No" columns, you may want to consider upgrading your **document use** skills.

3. Complete **Section 2** to identify your training needs.

Section 1: Self-Assessment

This table contains statements for the document use self-assessment.

I can...	Yes	Somewhat	No
Locate specific information in a simple document.			
Use key words or visual characteristics (e.g., colour, shape) to find information in a document.			
Create short lists (e.g., to-do list).			
Use a checklist.			
Recognize common workplace symbols, icons, and/or signs.			
Understand a simple invoice.			
Understand product or packaging labels.			
Enter information into simple forms, charts, or tables.			
Locate multiple pieces of information in a document.			
Use headings and sub-headings to find information in a document.			
Understand and locate information on charts or graphs (such as line graphs, bar graphs, or pie charts).			
Create simple diagrams to explain how something works.			
Use a table of contents or index page to find specific topics.			
Interpret detailed maps to find or give directions.			
Plot information onto different types of graphs.			
Enter several pieces of information into various forms, charts, or tables.			
Locate multiple pieces of information in a complex document.			
Understand and use a variety of complex documents.			
Understand and use information from several documents to solve a problem.			
Interpret blueprints.			
Understand and use complex diagrams.			
Create documents such as tables, charts, graphs, or scale drawings.			

I can...	Yes	Somewhat	No
Compare and integrate information from other documents to complete forms.			
Enter multiple pieces of information into complex forms, charts, or tables.			
Total			

Section 2: Personal development

Completing this section will help you make informed training decisions.

- Look at the "Yes" column in Section 1 to identify your strengths, and record them below.
- Look at the "Somewhat" and/or "No" columns in Section 1 to identify the areas that you need to develop or strengthen, and record them below.

Document use strengths:

I am confident that I can…

E.g., understand and locate information on charts or graphs.

1. _____

2. _____

3. _____

Tip: Consider using your strengths to help a colleague, friend or family member improve their document use skills.

Areas for improvement:

I would like to improve my ability to…

E.g., plot information onto different types of graphs.

1. _____

2. _____

3. _____

Tip: When developing your training plan, focus on improving one or two abilities at a time.

Numeracy

Instructions:

1. Read each statement in **Section 1** and place a check mark in the column that **best** describes how well you can complete that task.

    Tip: Think about your work and life experiences as you consider each task.

2. Review your responses for each task. If you have checked five or more in the "Somewhat" and/or "No" columns, you may want to consider upgrading your **numeracy** skills.

3. Complete **Section 2** to identify your training needs.

Section 1: Self-Assessment

This table contains statements for the numeracy self-assessment.

I can...	Yes	Somewhat	No
Perform simple calculations such as addition and subtraction.			
Receive cash payments and make change.			
Calculate the cost of items on a bill.			
Take simple measurements (e.g., length, weight, temperature).			
Make comparisons (e.g., taller or shorter, heavier or lighter, greater than or less than).			
Record time using digital and standard clocks, watches, or timers.			
Estimate quantities (e.g., I need approximately 20 copies).			
Estimate measurements (e.g., it is approximately three feet wide).			
Perform calculations that require multiplication and/or division.			
Calculate percentages.			
Calculate the area of common shapes (e.g., square, triangle, circle).			
Create and balance budgets.			
Create and monitor schedules (e.g., staffing or project schedules).			
Perform measurement conversions (e.g., inches to centimetres, millilitres to litres).			
Calculate simple averages.			

I can…	Yes	Somewhat	No
Estimate the time required to complete specific tasks.			
Perform calculations that require multiple steps or operations.			
Calculate areas and volumes of irregular shapes.			
Measure curved and irregular lengths.			
Take precise measurements using specialized equipment.			
Analyze and compare statistical data.			
Compare similar products with differing cost structures to determine the best value.			
Manage complex budgets (e.g., preparing financial statements, forecasting materials).			
Make accurate estimates when information is limited.			
Total			

Section 2: Personal development

Completing this section will help you make informed training decisions.

▪ Look at the "Yes" column in Section 1 to identify your strengths, and record them below.

▪ Look at the "Somewhat" and/or "No" columns in Section 1 to identify the areas that you need to develop or strengthen, and record them below.

Numeracy strengths:

I am confident that I can…

E.g., receive cash payments and make change.

1. _____

2. _____

3. _____

Tip: Consider using your strengths to help a colleague, friend or family member improve their numeracy skills.

Areas for improvement:

I would like to improve my ability to…

E.g., calculate areas and volumes of irregular shapes.

1. _____

2. _____

3. _____

Tip: When developing your training plan, focus on improving one or two abilities at a time.

Continuous learning

Instructions:

1. Read each statement in **Section 1** and place a check mark in the column that **best** describes how well you can complete that task.

    Tip: Think about your work and life experiences as you consider each task.

2. Review your responses for each task. If you have checked five or more in the "Somewhat" and/or "No" columns, you may want to consider upgrading your **continuous learning** skills.

3. Complete **Section 2** to identify your training needs.

Section 1: Self-Assessment

This table contains statements for the continuous learning self-assessment.

I can...	Yes	Somewhat	No
Learn new things.			
Ask questions when I do not understand something.			
Ask for feedback and/or advice from more experienced co-workers.			
Identify learning or training programs that are available to me at work and in my community.			
Learn by observing more experienced co-workers.			
Find and use learning materials and/or resources (e.g., searching the Internet, reading articles).			
Seek out and participate in training courses.			
Identify and understand my skill strengths and the areas where I need improvement.			
Develop my own learning goals at work and in my personal life.			
Apply the lessons I have learned from past experiences to new situations.			
Try new ways of doing things.			

I can...	Yes	Somewhat	No
Use newly learned skills and knowledge to improve my work.			
Recognize my preferred learning style (e.g., learning by seeing, hearing or doing).			
Be responsible for my own learning.			
Maintain my skill levels by practising what I have learned.			
Total			

Section 2: Personal development

Completing this section will help you make informed training decisions.

- Look at the "Yes" column in Section 1 to identify your strengths, and record them below.
- Look at the "Somewhat" and/or "No" columns in Section 1 to identify the areas that you need to develop or strengthen, and record them below.

Continuous learning strengths:

I am confident that I can...

E.g., ask for feedback and advice from more experienced co-workers.

1. _____
2. _____
3. _____

Tip: Consider using your strengths to help a colleague, friend or family member improve their continuous learning skills.

Areas for improvement:

I would like to improve my ability to...

E.g., try new ways of doing things.

1. _____
2. _____
3. _____

Tip: When developing your training plan, focus on improving one or two abilities at a time.

Working with Others

Instructions:

1. Read each statement in **Section 1** and place a check mark in the column that **best** describes how well you can complete that task.

   Tip: Think about your work and life experiences as you consider each task.

2. Review your responses for each task. If you have checked five or more in the "Somewhat" and/or "No" columns, you may want to consider upgrading your **working with others** skills.

3. Complete **Section 2** to identify your training needs.

Section 1: Self-Assessment

This table contains statements for the working with others self-assessment.

Working independently

I can...	Yes	Somewhat	No
Work independently to complete my tasks.			
Work with limited direction or supervision.			
Budget my time to ensure I complete tasks on time.			
Schedule and coordinate my work with the work of others.			
Organize my work tasks within a set of priorities.			
Take initiative by doing what needs to be done before being asked.			
Focus on my work even when there are distractions around me (e.g., other colleagues, background noise).			
Review my work to make sure that it is free of errors and meets quality standards.			
Ask for help or advice from co-workers or my supervisor when it is required.			
Provide regular progress reports to my supervisor so that he/she is aware of the work I do.			
Total			

Working with a partner or a team

I can...	Yes	Somewhat	No
Work co-operatively with a partner or team to complete tasks.			
Coordinate my work with the work of my colleagues to complete group projects.			
Complete my assigned work on time so that team deadlines are met.			
Complete my fair share of tasks when working with a partner or team.			
Follow directions from my partner or team members as required.			
Give directions to my partner or team members as required.			
Participate in making group decisions by contributing my ideas and suggestions.			
Contribute to making decisions co-operatively and settling differences respectfully.			
Improve my work based on suggestions and advice I receive from my partner or other team members.			
Help build an open and trustworthy work environment by encouraging others to participate in team building activities.			
Total			

Working in a leadership role

I can...	Yes	Somewhat	No
Coach or mentor others (e.g., help others by sharing my experiences and offering guidance or advice).			
Lead by setting a good example for the people around me.			
Support and encourage others.			
Let people know when they are doing a good job.			
Make decisions that I feel others can respect.			
Provide constructive feedback to help others improve their work.			
Take the lead in coordinating my colleagues' tasks in a group project.			
Demonstrate passion and enthusiasm for the work I do.			

I can...	Yes	Somewhat	No
Encourage group interactions and maintain a positive atmosphere within my team.			
Support colleagues by taking the time to help others with their work.			
Total			

Section 2: Personal development

Completing this section will help you make informed training decisions.

- Look at the "Yes" column in Section 1 to identify your strengths, and record them below.
- Look at the "Somewhat" and/or "No" columns in Section 1 to identify the areas that you need to develop or strengthen, and record them below.

Working with others strengths:

I am confident that I can…

E.g., focus on my work, even when there are distractions around me.

1. _____
2. _____
3. _____

Tip: Consider using your strengths to help a colleague, friend or family member improve their working with others skills.

Areas for improvement:

I would like to improve my ability to…

E.g., take the lead on group projects.

1. _____
2. _____
3. _____

Tip: When developing your training plan, focus on improving one or two abilities at a time.

Thinking

Instructions:

1. Read each statement in **Section 1** and place a check mark in the column that **best** describes how well you can complete that task.

   Tip: Think about your work and life experiences as you consider each task.

2. Review your responses for each task. If you have checked five or more in the "Somewhat" and/or "No" columns, you may want to consider upgrading your **thinking** skills.

3. Complete **Section 2** to identify your training needs.

Section 1: Self-Assessment

This table contains statements for the thinking self-assessment.

Problem solving

I can...	Yes	Somewhat	No
Identify the cause of a problem when I have all the necessary information given to me.			
Follow existing procedures or instructions to identify solutions to a problem (e.g., the steps for fixing a broken machine).			
Find information from a variety of sources (such as equipment manuals, policies and procedures) that will help me understand the problem and identify solutions.			
Use problem solving experiences I had in the past to help me identify solutions to current problems.			
Recognize key facts and issues related to a problem (e.g., identify answers to who, what, when, where, why and how).			
Identify and evaluate the pros and cons of each potential solution.			
Make adjustments to existing workplace procedures to help solve a problem (set procedures may not address every type of problem).			
Evaluate how well a solution worked.			
Total			

Decision making

I can...	Yes	Somewhat	No
Make decisions when following existing procedures or policies.			
Use my knowledge and past experiences to help me make decisions.			
Consider all the relevant information available before making a decision.			
Decide which of several options is most appropriate.			
Explain why I chose a particular decision.			
Total			

Job task planning and organizing

I can...	Yes	Somewhat	No
Complete tasks by their level of importance that have been organized for me.			
Complete tasks by their level of importance that I have organized on my own.			
Use tools such as calendars, agendas and to-do lists to help me organize my tasks.			
Coordinate my work with the work of my co-workers (e.g., make a schedule for using a shared piece of equipment).			
Deal with interruptions so that they do not interfere with my work schedule.			
Total			

Section 2: Personal development

Completing this section will help you make informed training decisions.

▪ Look at the "Yes" column in Section 1 to identify your strengths, and record them below.

▪ Look at the "Somewhat" and/or "No" columns in Section 1 to identify the areas that you need to develop or strengthen, and record them below.

Thinking strengths:

I am confident that I can…

E.g., apply past experiences to new problems or decisions.

1. _____

2. _____

3. _____

Tip: Consider using your strengths to help a colleague, friend or family member improve their thinking skills.

Areas for improvement:

I would like to improve my ability to…

E.g., make sure that minor interruptions do not interfere with my work plans.

1. _____

2. _____

3. _____

Tip: When developing your training plan, focus on improving one or two abilities at a time.

Appendix E: Ontario Skills Passport – Sample questions from general assessments

Read the task and answer **Yes** or **No**.

Reading

Do I think I can do this?

Read material safety data sheets (MSDS) to learn about the properties and uses of chemicals. (3)

Yes >>   No >>

Writing

Do I think I can do this?

Write reports to supervisors recommending the purchase of specific equipment or suggesting procedural changes. Use the reports to support and justify decisions. (3)

Yes >>   No >>

Document Use

Do I think I can do this?

Interpret data from a variety of graphs. For example, examine line graphs of hours worked and costs incurred. Interpret the graphs to identify patterns such as increased labour costs. (3)

Yes >>   No >>

Digital Technology

Do I think I can do this?

Click on an icon to open an app on a Smartphone. For example, click on an app to view the weather forecast before heading to work. (1)

Yes >>   No >>

Oral Communication

Do I think I can do this?

Coach and mentor junior trainees and offer guidance throughout projects. Provide constructive criticism, build trainees' technical knowledge and offer recommendations for improvement. (2)

Yes >>   No >>

Money Math

Do I think I can do this?

Make small cash purchases for supplies and services. For example, purchase coffee and snacks for meetings. (1)

Yes >>   No >>

Scheduling or Budgeting and Accounting

Do I think I can do this?

Adjust daily schedules to accommodate unforeseen circumstances such as emergencies or crises. (2)

Yes >>   No >>

Measurement and Calculation

Do I think I can do this?

Calculate paint quantities required, using wall surface area and paint coverage rates. (3)

Yes >>   No >>

Data Analysis

Do I think I can do this?

Review test results to identify potential errors by comparing the data to standard ranges for a particular type of test. (1)

Yes >>   No >>

Numerical Estimation

Do I think I can do this?

Estimate the percentage of wear of auto brake shoes in order to inform customers. (2)

Yes >>   No >>

Job Task Planning and Organizing

Do I think I can do this?

You work in a dynamic environment that requires you to respond to different challenges every day. For example, as a director, you direct the work of actors, dancers, production workers and various technicians. You prepare for new projects and manage current projects at the same time. You schedule time to attend to administrative tasks, plan and monitor production schedules and meet with collaborators. You adjust your work plans to accommodate events such as bad weather, faulty equipment and illnesses. You may participate in creating artistic visions and operational plans for your organization. (4)

Yes >>   No >>

Decision Making

Do I think I can do this?

Make decisions about staff training. Consider training costs, time requirements and long-term benefits. (3)

Yes >>   No >>

Problem Solving

Do I think I can do this?

You find that a machine is broken or not working properly. You follow the usual procedures to ask that the machine be fixed or replaced. (1)

Yes >>   No >>

Finding Information

Do I think I can do this?

Find information about project work sites by speaking with co-workers, reading historical information in company files, viewing maps and searching government websites. (3)

Yes >>   No >>

Critical Thinking

Do I think I can do this?

Scientists assess the impact of disasters caused by natural or human forces on ecosystems including plant, animal and human life. They identify and evaluate immediate, short-term and long-term impact of disasters. (4)

Yes >>   No >>

Figure 31   OSP Sample Self-Assessment – Teamwork

Read the task and answer by choosing the option that best describes what you do.

Teamwork

I work co-operatively with my supervisor and co-workers to get the job done.

Always   Usually   Sometimes   Never   Not Applicable

I listen respectfully to the ideas of others, even if I don't agree with them.

Always   Usually   Sometimes   Never   Not Applicable

I respect team members who have different opinions and values.

Always   Usually   Sometimes   Never   Not Applicable

I respect the feelings of others when I give feedback.

Always   Usually   Sometimes   Never   Not Applicable

I take on my fair share of work when I work in a team.

Always   Usually   Sometimes   Never   Not Applicable

Appendix F: OLES self-assessments for the trades

Instructions:

1. Read each statement and place a check mark in the column that best describes your ability to complete the task. Think about work and life experiences when considering your response to each statement.

2. Review your responses. A skill with five or more check marks in the Somewhat and/or No columns indicates an area you should consider improving.

3. Record your results in the Skills Summary section to gain a better understanding of your training needs.

Reading

Understanding materials written in sentences or paragraphs (e.g., reports, contracts and manuals).

I am confident in my ability to…	Yes	Somewhat	No
Read and follow directions on products or labels.			
Read and understand emails, memos or letters.			
Read a catalogue to learn basic product information.			
Read several short documents, such as newsletters, brochures or magazine articles to find information.			
Read and understand formal documents, such as service contracts and incident reports.			
Read and follow directions in equipment manuals, installation guides or work orders.			
Read handbooks and manuals to find information.			
Read policies or procedures when handling a problem at work.			
Read a variety of documents to compare information, such as product specifications.			
Read and interpret regulations to comply with standards.			
Total			

Document Use

Finding, understanding or entering information (e.g., text, symbols, numbers) in various types of documents, such as tables or forms.

I am **confident in my ability to**…	Yes	Somewhat	No
Understand signs or symbols, such as safety signs.			
Understand labels, such as shipping or hazardous materials labels (e.g., WHMIS).			
Create short lists, such as material lists.			
Find information in a document.			
Enter information into forms, such as order forms or building permits.			
Understand graphs, tables or charts, such as production reports or load charts.			
Enter information into graphs, tables or charts, such as work schedules or invoices.			
Create diagrams to explain how something works.			
Compare information from a variety of documents.			
Understand drawings or sketches, such as maps, schematics or blueprints.			
Create drawings or sketches, such as scale drawings.			
Total			

Numeracy

Using numbers and thinking in quantitative terms to complete tasks.

I am confident in my ability to…	Yes	Somewhat	No
Perform one-step calculations, such as addition, subtraction, multiplication or division.			
Perform multi-step calculations and calculations using percentages, fractions, decimals or ratios.			
Convert numbers from one unit of measurement to another (e.g., metric to imperial).			
Plan or monitor schedules, budgets or inventories.			
Take precise measurements, such as length or temperature.			

I am confident in my ability to…	Yes	Somewhat	No
Calculate the dimensions, area or volume of different shapes.			
Analyze or compare numerical data to identify trends or compile statistics.			
Estimate unknown values, such as time, distance, volume or quantity.			
Calculate ratios and proportions (e.g., determine actual measurements using scale drawings).			
Perform calculations using geometry (e.g., calculate slopes or elevation).			
Total			

Writing

Communicating by arranging words, numbers and symbols on paper or a computer screen.

I am confident in my ability to…	Yes	Somewhat	No
Record information, such as telephone messages or tasks to be completed.			
Use correct grammar and spelling.			
Write to inform or request information.			
Tailor writing for a specific audience, such as a foreperson or a supplier.			
Express my opinions in writing.			
Write brief notes, such as log entries or reminder notes.			
Write letters to clients or complete incident reports.			
Write notes to co-workers.			
Write business plans or proposals.			
Write documents using a template, such as contracts or financial reports.			
Total			

Oral Communication

Communicating by arranging words, numbers and symbols on paper or a computer screen.

I am confident in my ability to…	Yes	Somewhat	No
Talk to co-workers about a work-related issue.			
Talk to clients about services to be provided.			
Speak comfortably in different situations or to different groups of people.			
Communicate information clearly and concisely (e.g., explain a work-related issue to a supervisor).			
Understand information or questions that are presented orally (e.g., instructions for repairing a piece of equipment).			
Gather information by asking questions.			
Listen to others without interrupting.			
Restate in my own words information that is presented orally.			
Speak in front of a group of people (e.g., lead a discussion with about a work-related issue).			
Persuade others by speaking convincingly.			
Total			

Working with Others

Interacting with others to complete tasks.

I am confident in my ability to…	Yes	Somewhat	No
Work with limited direction or supervision.			
Work with others to schedule and coordinate job tasks.			
Ask for help when required.			
Complete my assigned work on time to ensure team deadlines are met.			
Give or follow recommendations or instructions.			
Recognize the strengths and weaknesses of other team members.			

I am confident in my ability to…	Yes	Somewhat	No
Use feedback from co-workers to improve my work.			
Give feedback to help others improve their work.			
Resolve conflicts when working with others.			
Take on a leadership role (e.g., mentor, advisor).			
Total			

Thinking

Finding and evaluating information to make informed decisions or to organize work.

I am confident in my ability to…	Yes	Somewhat	No
Recognize and identify problems.			
Use past experiences to help solve problems or make decisions.			
Identify several reasonable options to address a problem.			
Evaluate options and choose the best course of action when confronted with a problem or a decision.			
Make reasonable assumptions when information is unavailable.			
Find and use relevant information required to complete a task.			
Organize job tasks according to their level of priority.			
Memorize information required for different job tasks.			
Plan and arrange job tasks to meet deadlines.			
Evaluate the accuracy or credibility of information.			
Total			

Computer Use

Using computers and other forms of technology.

I am confident in my ability to…	Yes	Somewhat	No
Use company-specific technology, such as two-way radios or computer-controlled machinery.			
Perform basic interactions with a computer (e.g., log on, manage electronic files).			
Use word processing software to perform tasks (e.g., produce or format text).			
Use spreadsheet software to perform tasks (e.g., organize data, create graphs).			
Use databases to perform tasks (e.g., find or verify customer information).			
Use graphics software to create visual aids (e.g., create basic drawings of parts and fixtures).			
Use email to communicate (e.g., send attachments, email a group of people).			
Use company-specific software (e.g., financial).			
Use the Internet to find information (e.g., online supplier catalogue).			
Resolve basic technical difficulties (e.g., replace an ink cartridge in a printer).			
Total			

Continuous Learning

Participating in an ongoing process of improving skills and knowledge.

I am confident in my ability to…	Yes	Somewhat	No
Understand my skills strengths and areas for improvement.			
Develop a learning plan with guidance.			
Seek learning opportunities, materials and/or resources.			
Learn from past experiences and apply lessons learned to new situations.			
Try new ways of doing things.			
Learn from others (e.g., seek feedback from an experienced journeyperson).			
Take responsibility for my own learning.			

I am confident in my ability to…	Yes	Somewhat	No
Apply new skills and knowledge.			
Keep my skills up-to-date.			
Ask questions when information is unclear.			
Learn by reading and researching.			
Total			

Skills Summary

- Identify your essential skills strengths – skills with less than five check marks in the Somewhat and/or No columns.

- Identify areas for improvement – skills with five or more check marks in the Somewhat and/or No columns.

- Record your results in the space provided.

- Use your results to develop a training plan.

- Look at the "Somewhat" and/or "No" columns in Section 1 to identify the areas that you need to develop or strengthen, and record them below.

My essential skills strengths (e.g., reading)

1. _____

2. _____

3. _____

Areas for improvement (e.g., working with others)

1. _____

2. _____

3. _____

Appendix G: Ontario Skills Passport occupation-specific assessments

Figure 32   OSP Sample Self-Assessment **–** For Dental Assistants (NOC 3411)

Check off tasks you have done.

Reading

☐ Read instructions and other text entries on product labels and packaging, e.g. read product labels for mixing instructions and special handling requirements. (**1**)

☐ Read short notes to co-workers, e.g. read short notes from receptionists to learn about special requests and late arriving patients. (**1**)

☐ Read text entries in administrative and reporting forms, e.g. scan insurance company forms to learn about the cost coverage of specific procedures. (**2**)

☐ Read flyers, brochures and other promotional material to learn about promotions and new products, e.g. read about service options for cephalometric X-ray equipment in suppliers' brochures. (**2**)

☐ Read reference books, e.g. read oral pathology reference books to learn about diseases of the gums and teeth. (**3**)

☐ Read journals, magazines and website articles to stay current on industry trends and broaden your knowledge of techniques and materials. (**3**)

☐ Read manuals and guidelines to learn about methods and procedures for your work, e.g. read guidelines published by the Canadian Dental Association to learn about oral care for older adults. (**3**)

Writing

☐ Write short text entries in reporting and administrative forms, e.g. write the procedures completed on dental charts. (**1**)

☐ Write reminders and short notes to co-workers and colleagues, e.g. write short notes to inform specialists about new referrals. (**1**)

☐ Write instructions, e.g. write notes for patients outlining procedures that have been done and instructions for medication and other post-operative care. (**2**)

☐ Write letters to patients and their relatives on behalf of dentists, e.g. write a letter to a patient who did not show up for an appointment. (**2**)

Document Use

☐ Scan a variety of symbols and icons, e.g. scan Workplace Hazardous Materials Information System (WHMIS) symbols to identify potential biohazards. (**1**)

☐ Locate data, such as dates, sizes, codes and quantities, on labels, e.g. locate ingredient concentration levels on product labels. (**1**)

☐ Locate data in lists, tables and schedules, e.g. locate data, such as product names, identification numbers, classifications, quantities and costs, in suppliers' product lists. (**2**)

☐ Scan X-rays to determine their quality and usefulness. (**3**)

☐ Complete a variety of forms including inventory control forms, receipts, order forms, laboratory and X-ray requisition forms, insurance claims and dental charts by entering data, such as dates, times, quantities and costs. (**3**)

☐ Locate dimensions, angles and other data in various technical drawings, e.g. analyze assembly drawings to learn how to assemble and disassemble air-water syringes. (**3**)

Digital Technology

☐ Operate point-of-sale equipment, such as electronic cash registers, bar scanners, scales and touchscreens, to complete billings. (**1**)

☐ Operate electronic sanitization equipment, such as dry heat sterilizers and cage washers. (**1**)

☐ Use calculators and personal digital assistant (PDA) devices to complete numeracy-related tasks, such as calculating material requirements. (**1**)

☐ Use office equipment, such as printers, scanners, fax machines, copiers and postage meters, to perform clerical tasks. (**1**)

☐ Use browsers and search engines to locate information guidelines and bulletins from organizations, such as the Canadian Dental Association. (**2**)

☐ Use the Internet to access training courses and seminars offered by suppliers, employers and trainers. (**2**)

☐ Use word processing software to write letters. (**2**)

Oral Communication

☐ Talk to suppliers to learn about the availability of products and their delivery dates. (**1**)

☐ Reassure patients to reduce anxieties and to comfort those undergoing painful dental procedures. (**2**)

☐ Exchange information with co-workers, e.g. speak with co-workers to obtain information and coordinate activities. (**2**)

☐ Exchange information with dentists, e.g. listen to the dentist's directions before and during examinations and discuss patient treatments and conditions. (**2**)

☐ Talk to patients and their relatives, e.g. explain treatments and oral hygiene practices to young patients and their parents. (**2**)

Money Math

☐ Count cash and make change for cash payments, from patients, for exams and procedures not covered by dental insurance plans. (**1**)

☐ Calculate and verify invoice and receipt amounts, e.g. calculate amounts to be claimed from dental insurance plans. (**2**)

Scheduling or Budgeting and Accounting

☐ Compare options to get the best prices when ordering new supplies. (**1**)

☐ Create appointment schedules for dentists, frequently adjusting them because of cancellations and other unexpected events. (**2**)

☐ Calculate amounts for debit and credit transactions, accounts receivable and payable, bank reconciliations and summaries in general ledgers. (**3**)

Measurement and Calculation

☐ Measure pressures by reading numerical, colour-coded gauges when using sterilizers. (**1**)

☐ Calculate quantities of materials for mixtures, e.g. calculate the amount of water and plaster needed for a specific type of dental cast. (**2**)

Data Analysis

☐ Compare measurements of time, dimension and pressure to specifications, e.g. compare pressure readings of sterilization to specifications to determine when the equipment is ready for use. (**1**)

☐ Collect and analyze operational data, e.g. collect and analyze data on various types of dental procedures to calculate the average time spent on each type. (**2**)

Numerical Estimation

☐ Estimate time needed to perform job duties using past experience as a guide, e.g. estimate the time needed for an appointment by assessing the nature and complexity of dental procedures to be performed. (**1**)

Job Task Planning and Organizing

☐ Dental assistants organize their workday according to the scheduled appointments. They assess what preparation is required for each patient, sometimes helping several patients at the same time. (**2**)

Decision Making

☐ Select patients to fill cancelled appointments. Use first-come, first-served rule or assign priorities based on the severity of dental health conditions. (**1**)

☐ Select methods to organize materials and office supplies. (**1**)

☐ Decide the order of tasks and the priorities, e.g. decide the order in which to prepare patients for dental treatments. (**1**)

☐ Select suppliers and the supplies and materials to purchase. Take into account factors, such as budgets and anticipated needs. (**2**)

Problem Solving

☐ Encounter patients with special needs, e.g. serve those in wheelchairs who cannot lift themselves into the dentist's chair. Address their needs, such as by lifting them into the chair with the help of another staff member. (**1**)

☐ Fall behind schedule. Try speeding up procedures, rescheduling procedures or reducing the time dentists spend talking to patients. (**2**)

☐ Encounter nervous patients who are fearful of dental procedures. Speak using a reassuring tone of voice and provide information to reduce fears and anxieties. (**2**)

Finding Information

☐ Locate information about patients by asking them questions and by reading charts for information about their medical condition and special needs. (**1**)

☐ Consult catalogues and co-workers to find out about new products and to compare products or suppliers. (**2**)

☐ Consult oral pathology reference books, co-workers and dentists for information about diseases of the gums and teeth. (**2**)

☐ Locate specifications, such as application times, colour codes and ingredients, by scanning product labels, packaging and information posted on manufacturers' websites and by talking with suppliers and co-workers. (**2**)

Critical Thinking

☐ Judge the performance of products, such as fluoride treatments and moulding compounds. (**1**)

☐ Judge the condition and cleanliness of tools, such as probes and mouth mirrors. (**1**)

☐ Evaluate the performance of equipment, such as saliva ejectors and evacuators. (**1**)

☐ Evaluate the usefulness of X-rays and teeth moulds. (**2**)

Appendix H: Assessment from the New Brunswick cluster training project

Customer Relations

1. Please indicate how much you agree with each statement by checking the appropriate box.

	Strongly disagree 1	Disagree 2	Neither agree nor disagree 3	Agree 4	Strongly agree 5	Not applicable to my job 6
I have all the skills needed to do my job well in the following areas: Understanding how my daily tasks contribute to meeting customer expectations.	☐	☐	☐	☐	☐	☐
Understanding how the way I interact with customers contributes to business success.	☐	☐	☐	☐	☐	☐
Having a positive attitude when dealing with customers (e.g., being willing to adapt and be flexible in a difficult situation).	☐	☐	☐	☐	☐	☐
Taking initiative when dealing with customers (e.g., making decisions independently; doing things before being asked).	☐	☐	☐	☐	☐	☐
Listening to and understanding customer needs and requests.	☐	☐	☐	☐	☐	☐
Speaking clearly and appropriately to address customer needs or requests.	☐	☐	☐	☐	☐	☐
Recognizing and responding to non-verbal expressions of customer dissatisfaction or discomfort.	☐	☐	☐	☐	☐	☐
Asking clarifying questions to confirm information received from customers (e.g., credit card information, reservation details, etc.).	☐	☐	☐	☐	☐	☐
Showing interest with non-verbal cues (e.g., eye contact, stopping what you're doing) when talking to customers.	☐	☐	☐	☐	☐	☐

	Strongly disagree 1	Disagree 2	Neither agree nor disagree 3	Agree 4	Strongly agree 5	Not applicable to my job 6
Responding to email requests or enquiries.	☐	☐	☐	☐	☐	☐
Listening to and understanding customer complaints.	☐	☐	☐	☐	☐	☐
Asking clarifying questions to confirm customer concerns.	☐	☐	☐	☐	☐	☐
Negotiating and conflict resolution (e.g., de-escalating conflict with customers).	☐	☐	☐	☐	☐	☐
Exploring options to resolve complaints, and deciding on a solution.	☐	☐	☐	☐	☐	☐
Documenting complaints and resolutions according to standard procedures.	☐	☐	☐	☐	☐	☐

2. **In general how often do you feel stressed on the job when dealing with customers?**

- ☐ Rarely
- ☐ Sometimes
- ☐ Regularly
- ☐ Almost every day
- ☐ Not applicable to my job (I don't deal with customers directly)

B. Service Excellence

3. **Please indicate how much you agree with each statement by checking the appropriate box.**

	Strongly disagree 1	Disagree 2	Neither agree nor disagree 3	Agree 4	Strongly agree 5	Not applicable to my job 6
I have all the skills needed to do my job well in the following areas:						
Knowing how and where to find information to solve problems that come up on the job.	☐	☐	☐	☐	☐	☐
Collaborating with co-workers to solve problems.	☐	☐	☐	☐	☐	☐
Dealing with tool and equipment problems.	☐	☐	☐	☐	☐	☐

	Strongly disagree 1	Disagree 2	Neither agree nor disagree 3	Agree 4	Strongly agree 5	Not applicable to my job 6
Going 'above and beyond' basic service requirements to create experiences that exceed customer expectations.	☐	☐	☐	☐	☐	☐
Providing service that encourages first-time customers to return.	☐	☐	☐	☐	☐	☐
Knowing and being able to talk about my business's services and facilities.	☐	☐	☐	☐	☐	☐
Knowing and being able to talk about the community (e.g., restaurants, cultural activities, transportation options) in which my business is located.	☐	☐	☐	☐	☐	☐

4. **In general how often do you feel stressed on the job while trying to maintain service excellence?**

- ☐ Rarely
- ☐ Sometimes
- ☐ Regularly
- ☐ Almost every day

C. Productivity

5. **Please indicate how much you agree with each statement by checking the appropriate box.**

	Strongly disagree 1	Disagree 2	Neither agree nor disagree 3	Agree 4	Strongly agree 5	Not applicable to my job 6
I have all the skills needed to do my job well in the following areas:						
Planning and organizing job tasks to complete high-priority tasks on time.	☐	☐	☐	☐	☐	☐
Reducing inefficiency (e.g., spending too much time on low-priority tasks).	☐	☐	☐	☐	☐	☐
Performing key job tasks accurately and being able to spot errors.	☐	☐	☐	☐	☐	☐

	Strongly disagree 1	Disagree 2	Neither agree nor disagree 3	Agree 4	Strongly agree 5	Not applicable to my job 6
Coping with distractions while performing key job tasks.	☐	☐	☐	☐	☐	☐
Completing job-related documents accurately.	☐	☐	☐	☐	☐	☐

6. **In general how often do you feel stressed on the job because of concerns about your productivity?**

   ☐ Rarely
   ☐ Sometimes
   ☐ Regularly
   ☐ Almost every day

D. Working with colleagues

7. **Please indicate how much you agree with each statement by checking the appropriate box.**

	Strongly disagree 1	Disagree 2	Neither agree nor disagree 3	Agree 4	Strongly agree 5	Not applicable to my job 6
I have all the skills needed to do my job well in the following areas:						
Understanding the roles and responsibilities of my co-workers.	☐	☐	☐	☐	☐	☐
Providing feedback to, and accepting feedback from, my co-workers.	☐	☐	☐	☐	☐	☐
Offering my input, opinions, and ideas.	☐	☐	☐	☐	☐	☐
Understanding my role as a team member, and contributing to team success.	☐	☐	☐	☐	☐	☐
Asking for help when I need it.	☐	☐	☐	☐	☐	☐
Identifying strategies to resolve conflicts with my coworkers.	☐	☐	☐	☐	☐	☐
Understanding my responsibilities and being accountable for my work.	☐	☐	☐	☐	☐	☐

8. **In general how often do you feel stressed on the job as a result of dealing with co-workers?**

   ☐ Rarely
   ☐ Sometimes
   ☐ Regularly
   ☐ Almost every day

E. Health and safety

9. **Are you certified by the National Food Safety Training Program, offered by the Tourism Industry Association of New Brunswick (TIANB)?**

   ☐ Yes
   ☐ No, but I have completed another food safety training course (please specify) _____
   ☐ No, I have not completed a food safety training course
   ☐ Not applicable to my job (SKIP TO Q31)

10. **How often do you feel confident when applying food safety standards as part of your job?**

    ☐ Rarely
    ☐ Sometimes
    ☐ Most of the time
    ☐ Every time

11. **Did you complete the Responsible Beverage – It's Good Business course, offered by the Tourism Industry Association of New Brunswick (TIANB)?**

    ☐ Yes
    ☐ No, but I have completed another responsible beverage training course (please specify)
    _____
    ☐ No, I have not completed a responsible beverage training course
    ☐ Not applicable to my job (SKIP TO Q33)

12. **How often do you feel confident when applying responsible beverage standards as part of your job?**

    ☐ Rarely
    ☐ Sometimes
    ☐ Most of the time
    ☐ Every time

Appendix I: SRDC key informant interview protocol

Preamble

The Office of Literacy and Essential Skills (OLES) has retained the services of the Social Research and Demonstration Corporation to conduct a comprehensive review of measurement options for assessing outcomes of Literacy and Essential Skills (LES) initiatives and to develop a framework to support a broader performance measurement strategy.

As you know, OLES has identified nine workplace Essential Skills, including core literacy and numeracy skills (such as Reading, Writing, Document Use, Numeracy, and Digital skills), and four soft skills that are receiving increasing attention in recent research, but for which objectives measures of performance are not widely available: Oral Communication, Critical Thinking, Working with Others, and Continuous learning.

To complement our review of measurement options, we would like to learn a bit more about the resources you have developed or used to assess essential skills, inside the classroom and in the workplace. For this conversation, we would like to focus on the four soft skills, though any pertinent information about tools used to assess LES are welcome.

Questionnaire

1. Can you speak about the process of developing these tools?

    o   How involved is industry in the development of the tools?

    o   Do you develop generic tools that can be applied in various contexts?

2. For which of the following four soft essential skills have you developed (or are you currently developing) tools to assess ability:

    a. Oral Communication

    b. Critical Thinking,

    c. Working with Others, and

    d. Continuous learning

[ASK ALL QUESTIONS BELOW FOR EACH SKILL/MEASURE INDICATED ABOVE]

   (a) What assessment tool have you developed to measure them? What is the name of the assessment tool?

   (b) Are the measures ready to be used in the field?

      a. If not, at what stage of development are they?

      b. If yes, have they been used in the field? With whom (specific sector?) and where (in the classroom, in the workplace?)

(c) Do they include subjective measures (e.g., self-assessments or employer assessments) or objective measures (e.g., standardized tests, observable skills; performance benchmarks)?

(d) What are some of the indicators used to measure these skills?

(e) Have you assessed the statistical properties of these measures (e.g., reliability or validity of these measures)?

(f) Were the tools you use to measure this skill developed for a specific sector or would they be applicable to multiple sectors?

(g) In what context have you used these tools (in the classroom, in the workplace)?

(h) At what stage would these tools be most useful (e.g., to assess learner needs, to measure training gains? Are they better suited to measure job performance in the workplace following training?)

[The skills that are not being assessed]

Are you planning to develop any tools measuring any of these soft skills in the future?

a. Which skill(s)?

b. For what context will you be developing these tools (e.g., the classroom, the workplace)?

c. Will you be focusing on a specific sector?

What are some of the challenges that you face in developing reliable and objective measures for these essential skills?

From your experience, are there certain indicators of soft skills that may be more difficult to measure than others?

SRDC25**SRSA**

Ottawa • Vancouver

Ottawa

55 Murray Street, Suite 400
Ottawa, Ontario
K1N 5M3

55, rue Murray, bureau 400
Ottawa (Ontario)
K1N 5M3

Vancouver

789 West Pender Street, Suite 440
Vancouver, British Columbia
V6C 1H2

789, rue Pender Ouest, bureau 440
Vancouver (Colombie-Britannique)
V6C 1H2

www.srdc.org   1 866 896 7732   info@srdc.org