# WorkKeys
# Workplace Documents
# Technical Bulletin

# Table of Contents

# Workplace Documents Technical Bulletin

## Section 1: General Description of the Workplace Documents Assessment

The Workplace Documents assessment constitutes one of three parts of the ACT WorkKeys® National Career Readiness Certificate® (NCRC). The Workplace Documents assessment was developed to be an updated workplace reading assessment replacing the Reading for Information assessment. In developing the Workplace Documents assessment, ACT utilized four significant sources of information: (a) knowledge gained through 23 years of administering the Reading for Information assessment, (b) knowledge gained through ACT's job profiling services that included thousands of jobs for which the Reading for Information assessment was identified as relevant, (c) input and feedback from a panel of external Subject Matter Experts (SMEs) with experience in workforce development, and (d) professional literature published over the past 50 years analyzing how students and workers read, interpret, and use written text in the workplace.

The Workplace Documents assessment, like all ACT WorkKeys assessments, is designed to measure relevant skills required for success in today's job market. WorkKeys assessment data is based on skills, not scores, and has the capacity to demonstrate readiness to learn and succeed in the workplace, rather than achievements or deficits.

ACT organized a Design Team, composed of individuals from various areas within ACT including Test Development Content, Measurement and Research, Industrial/Organizational Psychology, and Assessment Design, to review the Reading for Information assessment and develop its replacement, the Workplace Documents assessment. The Design Team was assisted by a group of external SMEs who provided feedback and recommendations on the assessment construct, proposed test blueprint, and sample items.

The Workplace Documents assessment was designed to measure the extent to which individuals can read and comprehend written documents in order to do a job. The documents—which include, but are not limited to, messages, emails, letters, directions, signs, notices, bulletins, policies, websites, contracts,

and regulations—are based on materials that reflect the actual reading demands of the workplace. The ability to read and comprehend written information is critical for workplace success. The Workplace Documents assessment measures skills that individuals use when they read workplace documents and use that information to make decisions and solve problems (Sabatini, 2015).

## 1.1 Reading Workplace Documents as a Foundational Workplace Skill

ACT defines foundational workplace skills as "the skills that are essential for conveying and receiving information that is vital to work-related training and success" (ACT, 2014). Job skills are different from foundational skills. Job skills are the skills required to perform a specific job. For example, registered nurses must develop multiple skills, including giving an injection to a patient, in order to successfully fulfill their job tasks. Foundational skills are more general than job skills. They are the skills that enable a person to learn specific job skills.

Foundational skills are often referred to as basic or academic skills that may be learned through formal schooling or on one's own. Foundational skills are often defined in terms of academic subjects, including reading, writing, mathematics, and science. These skills enable individuals to acquire job specific skills, communicate information with fellow workers, and engage in lifelong learning. Foundational skills are also described as "portable" in that, rather than being job specific, they can be applied across a wide variety of jobs and occupations.

## 1.2 Reading in the Classroom and the Workplace

To help delineate the construct of reading Workplace Documents, ACT reviewed the relevant literature on reading skills. In general, it was noted that reading instruction in the classroom does not always align with workplace needs. As highlighted below, reading and workplace research indicates that successful application of reading skills is situation-specific, with reading behaviors dictated by the reader's purpose and circumstances.

While electronic recordings can sometimes be substituted for live speech or demonstrations, the written word is still the most consistently available communication medium in the workplace. Employees who need to learn or review a procedure, verify previously encountered information, or find answers to job-related questions frequently do so by reading. Good reading skills can be the difference between success and failure in workplace tasks as diverse as gathering ideas for a presentation, safely using a power tool, or mixing a solution in a lab. Similarly, good communication skills are frequently cited in surveys of employers as one of the top requirements of today's jobs (National Network of Business and Industry Associations (NNBIA), 2015). The ability to comprehend and interpret workplace documents is a critical component of workplace communication.

In contrast to classroom reading selections, workplace reading materials are usually written by individuals more qualified by their content knowledge than their writing skills. While these materials may be intended to convey precise meaning, they are not always easy to understand. Such materials may be used to train employees on safety and work procedures, or to provide information on employee benefits such as insurance policies and retirement plans. Employees read many of these materials in order to make decisions about some immediate course of action. Other materials describe behaviors or circumstances that may be relevant to their jobs in a more general sense. In both cases, the employees' comprehension of the text should not be taken for granted.

According to Human Resources and Skills Development Canada (2004), "A great deal of workplace reading is 'reading to do,' with the reader taking various actions and assuming risks associated with error. The fact that the reader takes various actions as a result of reading materials changes the dynamics of reading considerably. That is why the person with hands-on experience to support the knowledge gained through reading is often the best equipped to carry out the work." Thus, one important difference between workplace and school reading is the degree to which individuals must directly apply information gathered from texts—often with serious consequences for themselves and their teams.

On the other hand, a primary function of reading in a school environment is to teach widely applicable literary skills. Not surprisingly, there are several foundational reading skills developed in primary and secondary schooling that transfer into workplace reading situations. Table 1.1 summarizes the essential differences and points of overlap between classroom reading and workplace reading. The differences are likely to be in the purposes for reading, the type of materials that are read, and the amount of help that readers can expect when they approach reading tasks.

**Table 1.1: Classroom Reading versus Workplace Reading**

| Reading in School | Reading in the Workplace | Points of Overlap |
| --- | --- | --- |
| **Texts** | | |
| **Typical Text Types** | | |
| • Literature (fiction and non-fiction)<br>• Informational textbooks and websites on different subjects<br>• Assignments and worksheets<br>• Informational websites<br>• Instruction and procedures | • A range of procedural and informational documents<br>  ◦ Instructions<br>  ◦ Notices<br>  ◦ Bulletins<br>  ◦ Policies<br>  ◦ Regulations<br>• Email messages, memos, and other communications<br>• Informational websites | • Informational texts and websites<br>• Instructions and procedures |
| **Authors** | | |
| • Literary authors<br>• Multiple or unspecified authors who contribute to a textbook | • Technical writers, content experts, specialists (e.g., lawyers), coworkers, and customers<br>• Multiple or unspecified authors who contribute to a document | • Multiple or unspecified authors who contribute to a document |
| **Text Complexity Features and Readability Levels** | | |
| • Texts selected and adjusted for grade level<br>• Theoretical, academic language with emphasis on concepts and symbolic meaning<br>• Largely prose organizational features and formatting (topic-focused paragraphs, sections, and chapters)<br>• Online texts with hyperlinks and various navigation features | • A wide range of levels, related to specific features of task<br>• Technical, job-specific language with emphasis on concrete tasks<br>• A wide range of organizational features and formatting suited to specific task and purpose (e.g., mixtures of paragraphs, bullets/numbered lists, and other formatting elements)<br>• Online texts with hyperlinks and various navigation features | • Precise terminology<br>• Texts organized into paragraphs, sections, and chapters<br>• Interactive online texts |

Applying the above-discussed requirements of the workplace, ACT designed the Workplace Documents assessment to assess a wide range of skills related to reading and understanding workplace information, instructions, procedures, and policies. The action-oriented texts found in many workplaces differ from the explanatory and narrative texts on which most academic reading programs are based. In addition, unlike academic texts, which are usually organized to ease understanding and facilitate learning, workplace communication is not necessarily well written or written with ease of reading as a primary consideration. The reading selections in Workplace Documents are based on actual workplace materials representing a variety of occupations and workplace situations.

## 1.3 Workplace Documents—Assessment Claims

To validate test score interpretations and/or decisions is to review and evaluate the plausibility of the claims made regarding the test and its scores. Kane (2013) maintained that an argument-based approach to validation requires that the score-based claims be clearly articulated along with their associated inferences and assumptions. Validation therefore is a scientific process designed to evaluate the degree to which the analytic and empirical evidence supports the assessment claims along with the inferences and assumptions required to build the connections from examinee task performance to score-based interpretations and uses.

Based on the proposed uses of the Workplace Documents assessment, the Design Team defined three major claims focusing on workforce development concerns related to worker success and productivity at the individual level, business level, and regional level. Each claim and its underlying assumptions required for it to be plausible is outlined below.

**Claim #1:** U.S. examinees of high school or workforce age who demonstrate scores that reach at least a given level on the Workplace Documents assessment are more likely to successfully perform in more and higher levels of U.S. jobs (in the ACT job taxonomy) than examinees whose scores do not reach that level.

**Claim #1 Assumptions:**

1. The skill of reading workplace documents is a component of foundational workplace skills, and it is required for success in a large number of jobs (based on ACT's job profile database).

2. ACT has developed a professionally valid and appropriate definition of the reading workplace documents construct.

3. ACT's Workplace Documents assessment provides reliable and interpretable scores that reflect the construct. ACT's Workplace Documents assessment elicits observable evidence of the construct.

4. ACT has defined appropriate Workplace Documents Performance Level Descriptors (PLDs), and ACT has established standards (e.g., cut points) aligned to the PLDs.

5. Cut scores used to delineate each performance level have sufficient classification accuracy.

6. Businesses and employers are able to validly measure worker performance.

7. Scores on the Workplace Documents assessment are positively related to measures of employee performance, including productivity and turnover rates.

8. Examinees who score well on Workplace Documents are more likely to receive higher performance ratings and are more likely to have greater job success (defined as job retention and performance evaluations) than lower scoring examinees.

**Claim #2:** U.S. companies that hire U.S. examinees of high school or workforce age who demonstrate scores that reach at least a given level on the Workplace Documents assessment are more likely to achieve greater gains in productivity (for example, measured as increased output per day) from new employees than if the company had hired examinees whose scores do not reach that level.

**Claim #2 Assumptions:**

1. Claim #1 Assumptions 1–7

2. Employees who possess higher foundational workplace skills (as defined by ACT) are more likely to be productive and effective workers (as defined by supervisor evaluations) than employees who possess lower foundational workplace skills.

3. Having more productive workers leads to a business that is more effective and productive.

**Claim #3:** U.S. companies that hire U.S. examinees of high school or workforce age who demonstrate Workplace Documents scores that reach at least a given level are more likely to reduce turnover (retain those examinees for at least 6 months) than if the companies had hired examinees whose scores do not reach that level.

**Claim #3 Assumptions:**

1. Claim #1 Assumptions 1–7

2. Employees with higher foundational skill levels are less likely to be terminated in the first 6 months of employment than employees with lower foundational skill levels.

3. Employees with higher foundational skill levels are less likely to quit in the first 6 months of employment than employees with lower foundational skill levels.

4. Businesses that utilize scores from the Workplace Documents assessment as part of their hiring process will tend to experience less turnover than businesses who do not use the Workplace Documents assessment as part of their hiring process.

The three Workplace Documents claims focus on questions around examinee job success, improving worker productivity, and reducing employee turnover rates. Based on the claims, the critical stakeholders and intended test users are employers and hiring managers, state or regional workforce development officials, schools that prepare students to take jobs in the state or region, and examinees who are or will be seeking employment and career advancement.

Included as a part of all three claims are issues related to reading workplace documents as a component of foundational workplace skills, ACT's measure of reading workplace documents skills, and ACT's construct definition. These are critical assumptions that must be plausible to support the basis for test score interpretations and uses. (For information and data supporting the four underlying assumptions, please refer to Sections 7 and 8.)

# 1.4 Test Users and Stakeholders

The critical stakeholders are business employers, regional workforce development offices, schools that use the assessment as a measure of workforce readiness, and states or regions committed to developing their workforce. They are the individuals and groups who are invested in finding the right people for the right jobs.

Examinees. Individuals who take the Workplace Documents assessment are students and workers interested in demonstrating their foundational skill level in order to qualify as career ready, receive specific skill-related training, or qualify for a specific job. The examinee group includes individuals in high school through the adult working lifetime. High school students take the assessment to gain an understanding of their level of career readiness in critical skill areas and/or as a part of state accountability programs. Community college students take the assessment to demonstrate that they possess foundational skills and are ready to move forward for advanced training. College graduates take the assessment to demonstrate their level of career readiness as a means of differentiating themselves from other graduates. Working adults take the assessment to either qualify for a job or to demonstrate that they have the foundational skills needed for promotion or advanced training. In short, the examinee group includes high school students and adults who are either seeking employment or looking to advance in their field.

Stakeholders. Stakeholder groups include high schools and local school districts, state departments of education, community colleges, state and local workforce development departments, and employers.

High schools and local school districts administer the WorkKeys assessments in order to evaluate whether their curricular programs are enabling students to become career ready. In doing this, they are also providing their students the opportunity to earn a career ready certificate. State departments of education use the WorkKeys assessments as an accountability measure for evaluating the effectiveness of high schools and school districts in assisting their students to become career ready.

More specifically, the Workplace Documents assessment provides high schools and school districts with student data regarding the extent to which students have mastered facets of the curriculum associated with comprehending and interpreting written text materials in order to take an action or make a decision. The application of reading skills to workplace scenarios differentiates the Workplace Documents assessment from other standardized assessments of students' reading ability. The assessment provides the students the opportunity to demonstrate their reading mastery along with the skill of applying information gained through written text to solve work-related problems.

Community colleges utilize the Workplace Documents assessments in a variety of ways. Many community colleges use it as part of the process for determining acceptance into Career and Technical Education programs. Other community colleges use the assessment for program evaluation. Additionally, some community colleges use the assessment as a means of assisting their graduates in obtaining employment.

State and local workforce development offices utilize the assessment as a means of assisting unemployed or underemployed individuals in finding employment or better opportunities. The assessment provides a means for the workforce development office personnel to better understand the skill levels of individuals and to provide better guidance and assistance to them in finding employment.

Employers may use the Workplace Documents assessment, when coupled with a job profile analysis, to assist them in screening job applicants and finding sufficiently-qualified employees. A WorkKeys Job Profile allows the employer to understand the level of skill needed by a newly hired employee to successfully meet job expectations. Following the profile process, the employer may have job applicants take the assessment and then use their test scores as an additional piece of information to evaluate applicants.

# 1.5 Alignment to ACT's Holistic Framework

Building on research conducted over the last 50 years, ACT has developed its Holistic Framework (Camara, O'Connor, Mattern, & Hanson, 2015), which provides a more complete description of college and career readiness. The framework is organized into four broad domains: core academic skills, cross-cutting capabilities, behavioral skills, and education and career navigation skills.

1.  Core academic skills include the domain-specific knowledge and skills necessary to perform essential tasks in the core academic content areas of English language arts, mathematics, and science.

2.  Cross-cutting capabilities include the general knowledge and skills necessary to perform essential tasks across academic content areas. This includes technology and information literacy, collaborative problem solving, thinking and metacognition, and studying and learning.

3.  Behavioral skills include interpersonal, self-regulatory, and task-related behaviors important for adaptation to and successful performance in education and workplace settings.

4.  Education and career navigation skills include the personal characteristics, processes, and knowledge that influence individuals as they navigate their educational and career paths (e.g., make informed, personally relevant decisions; develop actionable, achievable plans).

The Workplace Documents assessment aligns primarily with the first broad domain of the Holistic Framework, which included domain-specific knowledge and skills necessary for performing essential tasks. The assessment uses authentic workplace documents and scenarios in order to determine an examinee's level of proficiency in reading workplace documents and applying the information within these documents to the types of tasks an employee would be expected to perform. The ability to use and interpret entire texts or parts of a text, summarize a text, locate key details, draw conclusions and inferences, and understand vocabulary used in context are foundational skills necessary in both academic and workplace settings. As such these skills are the focus of the Workplace Documents assessment and align this assessment to the skills defined in the Holistic Framework of education and work readiness.

# Section 2: Test Development

WorkKeys Workplace Documents is designed to assess the extent to which individuals can read and comprehend written documents in order to do a job. Through a review of the pertinent empirical and professional literature and through deliberations among team members, the Design Team determined that the Workplace Documents reading construct was defined through the interplay of two facets: Document Level Complexity and Reading Skills. Although each aspect is defined separately, collectively they interact to provide meaning and interpretability to test scores.

For the Workplace Documents reading construct, reading skill progressions are highly relevant. As a result, the team began by defining the characteristics of different levels of reading difficulty, and then by identifying the pertinent associated reading skills.

## 2.1 Document Level Complexity

Document level complexity refers to the text complexity of the reading documents examinees are required to read in order to respond to the items. The Design Team organized document level complexity into five levels. Document (text) complexity for the Workplace Documents assessment is defined by the document's word count, Flesch-Kincaid Reading Grade Level, clarity, amount of detail, and vocabulary level (including the use of technical terms, jargon, and acronyms). Both word count and the reading level may be overridden based on other criteria. For stimuli containing more than one document, all documents are combined to determine the word count and grade level. Because the contracts, legal, and multiple related document types are inherently more challenging than others, they are permitted only at the higher levels.

Table 2.1 provides the Workplace Documents complexity criteria along with the descriptor for each level.

**Table 2.1: Workplace Documents—Passage Level Complexity Descriptors**

| WD Document Criteria | Level 3 | Level 4 | Level 5 | Level 6 | Level 7 |
|---|---|---|---|---|---|
| Word Count | Range: 80–150 | Range: 100–200 | Range: 150–350 | Range: 200–450 | Range: 250–500 |
| Flesch-Kincaid Reading Grade Level* | 6 | 7.5 | 10 | 12 | 13 |
| How complex is the stimulus document? | Short with no extra information and simple sentences | Straightforward with some longer sentences; may contain conditional situations | Mostly clear and direct, but with multiple details; may have complex sentences and/or contain conditional situations | Somewhat complicated sentences, document may be long and/or complex and/or contain conditional situations | Complex sentences with many details; may cover uncommon topics and/or contain conditional situations |
| Is the information in the document clearly stated? | Yes | Yes, mostly | Not necessarily; may need to make inferences | No, information is often not explicit | No, pieces of information may be spread throughout documents and may be extraneous |
| How detailed is the document? | Not very; will include a small number of details | There are a number of details | There are many details and some may be extraneous | There are implied and/or extraneous details | There are many implied and extraneous details |
| How difficult is the vocabulary? | Common, familiar, not difficult | Not too difficult; common vocabulary with some advanced words | Unfamiliar words, professional jargon, and acronyms; may need to use context to determine correct meaning | Difficult words, professional jargon and technical terms; meanings may need to be determined from context | Advanced, unfamiliar, and/or uncommon words, technical terms, and professional jargon; meanings must be determined from context |
| Document Type | Informational, Instructional, Policy | Informational, Instructional, Policy | Informational, Instructional, Policy, Contracts, Legal, Multiple Related Documents | Informational, Instructional, Policy, Contracts, Legal, Multiple Related Documents | Informational, Instructional, Policy, Contracts, Legal, Multiple Related Documents |

*The Flesch-Kincaid Reading Grade Level is a quantitative measure of the level of readability.

Workplace Documents—Document Classification Consistency Study. ACT conducted a study to evaluate the Content Specialists' ability to consistently classify different reading passages into the five levels based on the criteria described in Table 2.1. The study asked four content specialists with substantial experience developing the Reading for Information assessment to discuss how they classified workplace documents and the merits of using the table to determine the level of such documents. Following the discussion, the four content specialists independently evaluated 20 reading passages and classified them into one of the five levels.

ACT utilized Generalizability Theory (Brennan, 2001) to analyze the consistency of the content specialists' categorizing. A graphics x rater design was modeled and used the GENOVA software program (Crick & Brennan, 2001) to analyze the ratings. The analysis provided a Generalizability Coefficient of 0.93, and a Phi Coefficient of 0.92. These consistency indices revealed that the four content specialists, using Table 2.1 along with their training, classified workplace documents in a consistent manner.

## 2.2 Workplace Documents Skill Domain Definitions

ACT's reading content specialists reviewed the original list of reading skills measured through the Reading for Information assessment. They determined that several of the defined skills overlapped and caused confusion in identifying the skill that aligns to the item. Consequently, the Design Team concluded that many of the old Reading for Information skill definitions were confusing and needed to be simplified.

In an effort to achieve greater clarity regarding the skill definitions, the content specialists reviewed the professional literature on reading and the workplace, and they asked the external SMEs for direction and insight. Through this work, they concluded that three primary skill domains exist in regards to reading workplace documents: comprehending written text, interpreting written text, and applying information and instructions derived from written text to workplace situations. As a result, the Design Team identified three primary reading-related workplace skills:

- • Identify Main Ideas and Details
- • Apply Instructions or Information
- • Identify Meanings and Definitions of Words or Phases

From these three primary skills, they defined a progression of reading subskills within each primary skill relevant to workplace applications. The workplace reading skills and subskills progression is presented in Table 2.2.

**Table 2.2: Workplace Documents Skills Subskills**

| | Skills and Subskills |
|---|---|
| **1.0** | **Identify Main Ideas and Details** |
| 1.1.a. | Identify the main idea |
| 1.1.b. | Identify the rationale behind an entire document or a section of a document<br>*Identify an underlying reason for a task or procedure. Often, "what is the main reason . . . ?"* |
| 1.2.a. | Identify specific details |
| 1.2.b. | Infer implied details<br>*The details needed to complete a task or procedure are not explicit at all; inferences need to be made to determine the necessary information.* |
| **2.0** | **Apply Instructions or Information** |
| 2.1 | Choose when to perform a step in a series of steps<br>*Often includes questions such as "What should you do first/next/last?"* |
| 2.2.a. | Apply information/instructions to a described situation<br>*Identify the necessary information/instructions to complete a task and correctly apply them to a situation described in the document—"You should . . ."* |
| 2.2.b. | Apply information/instructions to a situation not directly described or to a completely new situation<br>*Identify the necessary information/instructions to complete a task and correctly apply them to a situation that is not described in the document.* |
| 2.2.c. | Apply principles inferred from a passage to a situation not directly described or to a completely new situation<br>*Infer the reasons behind instructions/information described in the document and correctly apply them to a situation that is not described.* |
| **3.0** | **Identify Meanings and Definitions of Words and Phrases** |
| 3.1 | Infer the meaning of a word or phrase from context (not jargon or technical terms)<br>*Infer the correct meaning of words and phrases as they are used in a specific workplace scenario from the context of the document.* |
| 3.2.a. | Identify the meaning of an acronym, jargon, or a technical term<br>*Identify the meaning of words, phrases, acronyms, or jargon that have an exclusive meaning in a particular job or career cluster.* |
| 3.2.b. | Infer the meaning of an uncommon acronym, jargon, or a technical term from context<br>*Infer the meaning of words, phrases, acronyms, or jargon that have an exclusive meaning in a particular job or career cluster from the context of the document.* |

After having the external SMEs review the skills and subskills described in Table 2.2, the team concluded that the most highly relevant workplace reading skill was Apply Instructions and Information. In a workplace context, employers and supervisors are most concerned that workers not only are able to read and understand written texts, but, more importantly, that they understand how and when to apply

instructions and information contained in the documents. Being able to apply information appropriately and accurately is critical to being a successful worker, and in today's workplace much of the information is presented to workers in written documents.

## 2.3 Workplace Documents—Multiple Related Documents

The Workplace Documents assessment includes the five document types used on the Reading for Information assessment: instructions, informational, policies, legal, and contracts. Additionally, the Workplace Documents assessment includes multiple related documents. To reflect real-world situations that often require an individual to identify information from multiple documents, make connections and conclusions, and apply this information to accomplish tasks, multiple related documents was added to the assessment.

The definition of the multiple related documents passage type is that they

- consist of two or more documents that are related or cover a common topic, and

- have two or more authors

Examples of these multiple related documents may include:

- an email string

- two webpages on a similar topic

- a company policy followed by a question raised in an email or message by a client or customer

- a formal document followed by an informal document that elaborates or explains

The Workplace Documents Design Team considered it critical to include this type of a document in the assessment so that it accurately represented the type of reading content and context workers use on a daily basis. It provides examinees the opportunity to demonstrate the ability to read complex text materials, understand and apply differing perspectives, and utilize the information contained in these documents to complete workplace tasks. In many ways, this change not only represents a unique passage type, but it also reflects the assessment's emphasis on authentic and up-to-date reading passages and item tasks (Binkley, Erstad, Herman, Raizen, Ripley, Miller-Ricci, & Rumble, 2012; Sabatini, 2015).

## 2.4 Workplace Documents—Performance Level Descriptors (PLDs)

The Workplace Documents assessment construct is defined through a combination of the text complexity level of a reading passage and the skill elicited by the item. Based on the text complexity level and skill, the Design Team was able to define the Workplace Documents PLDs.

Level 3—Document types include informational, instructional, and policy-related materials.

Examinees scoring at Level 3 are able to read and comprehend relatively short workplace documents which contain no extra information. The document contains short sentences using common, everyday

workplace vocabulary. All the information in these documents is clearly and directly stated, and it contains a small number of details. In reading these documents, they are able to:

- Identify the main idea
- Identify specific details
- Choose when to perform a step in a series of short steps
- Apply information/instructions to a situation that is the same as the situation in the reading materials

Level 4—Document types include informational, instructional, and policy-related materials.

Examinees scoring at Level 4 have the skills defined at Level 3 and in addition are able to read and comprehend workplace documents written in straightforward sentences that use familiar vocabulary and the occasional use of conditionals and a few advanced words. In reading these documents, they are able to:

- Identify the main idea
- Identify specific details
- Use the reading materials to figure out the meanings of words that are not defined for them
- Choose when to perform a step in a series of steps
- Apply information/instructions to a situation that is the same as the situation in the reading materials
- Choose what to do when changing conditions call for a different action

Level 5—Document types include informational, instructional, policy-related, contractual, legal, and multiple related document materials.

Examinees scoring at Level 5 have the skills defined at Levels 3 and 4, and in addition are able to read and comprehend longer workplace documents written in more complex sentences that use more advanced vocabulary, including unfamiliar technical words, jargon, and acronyms. The information in Level 5 documents is generally stated directly, but specific details may be more difficult to find because of extraneous information. In reading these documents, they are able to:

- Identify specific details
- Infer the meaning of a word or phrase from context
- Apply information/instructions to a new situation that is similar to the one described in the document while considering changing conditions
- Apply information/instructions that include conditions to situations described in the document
- Identify the appropriate meaning of an acronym, jargon, or technical term defined in the document
- Apply technical terms and jargon to stated situations
- Make some inferences to accomplish a goal

Level 6—Document types include informational, instructional, policy-related, contractual, legal, and multiple related document materials.

Examinees scoring at Level 6 have the skills defined at Levels 3, 4, and 5, and in addition are able to read and comprehend longer workplace documents written in lengthy, complex sentences that use advanced vocabulary including unfamiliar words, jargon, and acronyms where the meaning is often implied. In reading these documents, they are able to:

- Infer implied details

- Infer the meaning of an acronym, jargon, or technical term from context

- Apply information/instructions to a situation not directly described or to a completely new situation

- Apply principles inferred in a passage to a situation not directly described or to a completely new situation

- Identify the rationale behind an entire document or a section of a document

Level 7—Document types include informational, instructional, policy-related, contractual, legal, and multiple related document materials.

Examinees scoring at Level 7 have the skills defined at Levels 3, 4, 5, and 6, and in addition are able to read and comprehend long workplace documents with many details and are written using lengthy, complex sentences that use advanced vocabulary (including esoteric words, jargon, and acronyms) where meanings must be inferred from context. In reading these documents, they are able to:

- Infer implied details

- Infer the meaning of an acronym, jargon, or technical term from context

- Apply information/instructions to a situation not directly described or to a completely new situation

- Apply principles inferred in a passage to a situation not directly described or to a completely new situation

- Identify the rationale behind an entire document or a section of a document

# 2.5 Designing Items to Elicit Examinee Evidence of Reading Workplace Documents

Workplace Documents uses multiple-choice items to measure examinees' proficiency in reading and comprehending workplace texts to gain information to apply in workplace situations. The construct of workplace reading skills measured by the assessment was defined by the Design Team and confirmed by external SMEs with backgrounds in business, industry, and education. To properly elicit evidence of the skills in the workplace reading domain, ACT follows an item-design model aligned with both evidence-centered assessment design (Mislevy, Steinberg, & Almond, 1999) and the *Standards for Educational and Psychological Testing* (American Educational Research Association (AERA), American Psychological Association (APA), & National Council for Measurement in Education (NCME), 2014).

## 2.5.1 Item Writing

Item writers qualify to write for the Workplace Documents assessment by completing item-writing training modules. The modules cover numerous aspects of developing quality multiple-choice items including creating passages that elicit evidence of the skill the item measures, writing effective distractors, employing realistic workplace contexts, and avoiding common item-writing errors. For workplace reading, the training also provides explicit direction in terms of acceptable workplace reading texts. Once an item writer has successfully completed all required training modules, he or she is given an item-writing assignment that details the number of items to be developed at specific levels and for specific skills. Once an item writer has demonstrated the ability to write high-quality items by successfully completing the assignment, he or she receive materials explaining item task models.

The task models provide item writers with the following information: (a) skill name, (b) skill description, (c) evidence statement, (d) item components, and (e) item exemplars. Additional requirements related to the items include:

- All items are linked to a workplace-oriented passage

- Workplace passages are written documents taken from workplace situations and scenarios

- Workplace passages are evaluated by the content team in terms of workplace realism

- Workplace passages are evaluated by the content team and classified into one of the five levels

- Workplace passages must be designed for one of the following purposes: (a) convey information to workers; (b) instruct workers on a procedure, process, or other activity; (c) convey a workplace policy; (d) convey contractual and/or legal information; and (e) convey information in multiple related documents written by two or more authors

- Multiple items should be developed for each workplace passage

- Each item is aligned to one of the skills defined as part of the construct

## 2.5.2 Item Review

After items have been developed, edited, and tentatively finalized by the Content Assessment team, they are submitted to external consultants with experience in reading workplace documents for content and fairness reviews. In the content review, they evaluate the passage and item in terms of

- whether the passage and item are appropriately aligned to the construct

- whether the context and the solution method are authentic and workplace relevant

- whether there is one, and only one, correct response for each item

Reviewers also evaluate the passages and items for fairness and cultural bias by considering how members of different demographic groups would respond to them. (ACT asks the item reviewer to evaluate the item from the perspective of men and women examinees, and from the perspective of African-American, Hispanic-American, and Asian-American examinees.) The reviewer is asked to comment on whether there is anything within the item that any group might find offensive. Also, the reviewer is asked to evaluate if each demographic group has equal access to, and opportunity to learn, the information and skills assessed.

For both the content and fairness reviews, item reviewers complete a questionnaire either approving the item as written or identifying specific concerns. ACT content specialists gather the information from the reviewers and determine how to appropriately address any concerns. Items are not classified as ready for pretesting until all relevant issues are resolved.

### 2.5.3 Item Pretesting

All Workplace Document items are pretested before they become operational. Newly developed or recently revised items are embedded in current forms of the Workplace Documents assessment. As a result, examinees respond to the pretest items as a part of their responses to the operational assessment.

ACT conducts statistical analyses to determine if each pretest item meets required statistical criteria. ACT analyzes the items using both classical and item response theory (IRT) statistics to evaluate their psychometric properties, including item difficulty and discrimination. If the pretest item meets the statistical criteria, it has passed pretesting. If it fails to meet the criteria, the Workplace Documents content team reviews it and considers whether it should be edited, modified, or removed from the pool. When an item is edited or modified, it receives a new item identifier and is pretested a second time.

To ensure item fairness, ACT compares item difficulty values based on group membership (item analysis is conducted comparing difficulty levels by gender and ethnic status) and performs Differential Item Functioning (DIF) evaluations (Holland & Wainer, 1993). Items that are flagged through the DIF evaluations are sent to the Workplace Documents content team for review. The content team determines whether the flagged item should remain as it currently is, be revised and returned to pretesting, or be removed from the pool.

## Section 3: Test Specifications

An assessment's specifications are developed by first determining the assessment's claims and score interpretations, and then by articulating the set of behaviors that need to be elicited through the test content to provide evidence to support the claims. In articulating the set of behaviors, the team evaluates the degree to which examinee responses to the item content provide support for the assessment's claims and score interpretations. Item and test content must elicit examinee behaviors that are aligned to the construct and that provide evidence supporting score interpretations (Kane, 2013; Messick, 1989).

The Workplace Documents Design Team utilized a variety of reputable source materials to identify relevant content that constitutes a measure of workplace reading. Over the past 25 years, through its job profiling services, ACT has gathered information related to workplace reading texts, tasks, and skills from the manufacturing, health care, construction, transportation, financial, and sales sectors. The Workplace Documents team reviewed these findings and used the information to determine what types of reading materials should be included and which skills were most frequently required. To further support content-related decisions, the team reviewed professional literature around workplace reading (Binkley, et al., 2012; Smith, Mikulecky, Kibby, Dreher, & Dole, 2000) and workplace competency models (NNBIA, 2014). Lastly, the team consulted with a group of external SMEs to obtain their perspective on workplace reading texts and skills.

# 3.1 Content Relevance and Representativeness

Test specifications must be carefully defined to ensure that the assessment tasks are construct relevant and representative of the domain purported to be measured (Messick, 1989; Mislevy, Steinberg, & Almond, 1999). In the context of Workplace Documents, construct relevance requires not only that the examinee demonstrate the ability to read and comprehend a workplace document, but that he or she also demonstrates the ability to apply the information conveyed by the document to complete a job task. Because WorkKeys assessments are designed to measure skills that are widely applicable to a large number of jobs, construct representativeness refers to a range of reading passages and the various reading skills needed in the workplace. To illustrate, reading materials must represent the full range of job sectors, from manufacturing to construction to office work. The reading materials must also represent appropriate ranges of difficulty, from straightforward easy-to-read texts to more complicated and nuanced writings.

A second purpose of the test specifications involves the development of alternate forms. The size of the WorkKeys test population combined with the need for security and fairness necessitates the construction of alternate forms of Workplace Documents. In developing alternate forms, ACT requires that all forms must meet Lord's (1980) equity property. Lord's equity property states, from the test taker's perspective, it must be a matter of score indifference whether he or she is administered Form A or Form B of an assessment. To achieve alternate forms that meet the equity property, the content representativeness of each form must be identical (Kolen & Brennan, 2004).

By carefully defining the content specifications, ACT accomplishes two critical assessment goals:

1. Content is construct relevant and representative.

2. Content representation is identical across alternate forms.

# 3.2 Workplace Documents—Test Blueprint

ACT developed detailed blueprints defining the content attributes of each test item. The content specifications were developed by clearly specifying the attributes of a reading passage at each of the five levels (see Section 2: Test Development). They were further defined by specifying the workplace reading skills and subskills. Within the test specifications table, each subskill was evaluated and aligned to one or more levels. Following the alignment of skill subskills, weights were determined based on the overall importance of the subskill to the construct of workplace reading (Allen & Yen, 2002).

The Workplace Documents construct was based on two critical facets:

• Reading Complexity Level of the Passage

• Reading Skill elicited by the Item

The five Reading Complexity Levels were defined by the passage's word count, reading level, clarity, amount of detail, and vocabulary level (see Table 2.1). ACT content specialists evaluated each passage and, based on these factors, determined the appropriate level.

Workplace reading skills were divided into three primary skills: identify main ideas and details, apply instructions or information, and identify meanings and definitions of words or phrases. After analyzing the professional literature on workplace reading and data from ACT's job profiling, ACT determined the

importance of the skills to overall workplace success. As illustrated in Table 3.1, the number of items assessing each skill reflect its relative importance, with identifying meanings and definitions being least important and applying instructions or information being most important.

Five critical document types were identified as relevant to workplace reading: informational, instructional, policy, legal, and multiple related. At the lowest two levels (Level 3 and Level 4), only informational, instructional, and policy documents were considered relevant for workplace reading. At the intermediate level (Level 5), the most relevant documents were informational, instructional, and policy, although Level 5 passages may include a legal or multiple related document. At the two highest levels, all five document types were considered relevant.

Tables 3.1, 3.2, and 3.3 present the Workplace Documents test specifications. The test specifications provide a blueprint for form development and also represent the relative importance of the reading skills and subskills in the workplace.

**Table 3.1: Skill Domain Item Distribution by Level**

| Domain | Number per Level | | | | | |
|---|---|---|---|---|---|---|
| | Level 3 | Level 4 | Level 5 | Level 6 | Level 7 | Total |
| 1.0 Identify the Main Idea and Details | 2 | 3 | 1 | 2 | 2 | 10 |
| 2.0 Apply Instructions or Information | 2 | 4 | 4 | 3 | 2 | 15 |
| 3.0 Identify Meanings and Definitions of Words and Phrases | 0 | 1 | 2 | 1 | 1 | 5 |
| | 4 | 8 | 7 | 6 | 5 | 30 |

**Table 3.2: Skill Subdomain Item Distribution by Level**

| Subdomain | Number per Level | | | | | |
| --- | --- | --- | --- | --- | --- | --- |
| | Level 3 | Level 4 | Level 5 | Level 6 | Level 7 | Total |
| 1.1.a Identify the Main Idea | 1 | 1 | 0 | 0 | 0 | 2 |
| 1.1.b Identify the Rationale Behind an Entire Document or a Section of a Document | 0 | 0 | 0 | 1 | 1 | 2 |
| 1.2.a Identify Specific Details | 1 | 2 | 1 | 0 | 0 | 4 |
| 1.2.b Infer Implied Details | 0 | 0 | 0 | 1 | 1 | 2 |
| 2.1. Choose When to Perform a Step in a Series of Steps | 1 | 1 | 0 | 0 | 0 | 2 |
| 2.2.a Apply Information/Instruction to a Described Situation | 1 | 3 | 2 | 0 | 0 | 6 |
| 2.2.b Apply Information/Instructions to a Situation Not Directly Described or to a Completely New Situation | 0 | 0 | 2 | 1 | 0 | 3 |
| 2.2.c Apply Principles Inferred from a Passage to a Situation Not Directly Described or to a Completely New Situation | 0 | 0 | 0 | 2 | 2 | 4 |
| 3.1 Infer the Meaning of a Word or Phrase from Context (Nonprofessional) | 0 | 1 | 1 | 0 | 0 | 2 |
| 3.2.a Identify the Meaning of an Acronym, Jargon, or Technical Term Defined in a Document | 0 | 0 | 1 | 0 | 0 | 1 |
| 3.2.b Infer the Meaning of an Acronym, Jargon, or Technical Term from Context | 0 | 0 | 0 | 1 | 1 | 2 |
| Total | 4 | 8 | 7 | 6 | 5 | 30 |

**Table 3.3: Number of Passages for each Document Type**

| Document Type | Number of Passages per Form | Max Number of Passages (includes pretest) |
| --- | --- | --- |
| Instructional (INS) | 3–5 | 5 |
| Informational (INF) | 2–4 | 4 |
| Policy (POL) | 2–4 | 4 |
| Legal (LEG) | 1–3 | 3 |
| Multiple Related (MUL) | 1–3 | 3 |

Each form of the Workplace Documents assessment is built to conform to the test specifications defined in Tables 3.1, 3.2, and 3.3. ACT's test development and psychometric staff members thoroughly review each form to ensure that it meets the specifications, and that each form is parallel in terms of content to all other Workplace Documents forms.

# Section 4: Test Administration

The *ACT WorkKeys Administration Manual—Paper Testing* and *ACT WorkKeys Administration Manual—Online Testing* provide direction in the administration of the WorkKeys assessments including timing instructions. It is important that all staff involved in the administration of WorkKeys assessments follow the instructions as provided by ACT to appropriately measure the skills and abilities of the individuals completing the assessments.

## 4.1 Standard Test Administrative Procedures

Included in the two manuals are detailed directions for securing materials and administering the assessments in a standardized manner. The following actions violate ACT policies and procedures for delivering WorkKeys assessments:

- accessing or obtaining a test booklet or test questions prior to the test for any reason (An exception is provided for American Sign Language and Signing Exact English interpreters assisting examinees)
- photocopying, making an electronic copy, or keeping a personal copy of the test or of any test items
- taking notes about test questions or any paraphrase of test questions to aid in preparing examinees for testing
- aiding or assisting an examinee with a response or answer to a secure test item, including providing formulas
- rephrasing test questions for examinees
- creating an answer key or "crib sheet" of answers to test questions
- editing or changing examinee answers after completion of the test, with or without the examinee's permission
- allowing examinees to test in an unsupervised setting
- leaving test materials in an unsecured place or unattended
- failing to properly report and document incidents of prohibited behavior involving examinees, staff, or others
- allowing examinees to test longer than the permitted time
- failing to return and account for all testing materials after the testing session has ended

## 4.2 Test Administration Personnel

ACT identifies a Test Center Coordinator to select and reserve appropriate rooms required for testing, select and train qualified personnel to administer the assessments, order and ensure the security of all test materials, and coordinate all activities on test days. The Test Center Coordinator should have a back-up coordinator who is responsible for carrying out activities should the Test Coordinator not be able to do so.

ACT provides copies of the administrative manuals, which every staff member is expected to read, understand, and follow. The manual is proprietary information and is copyrighted by ACT. It is to be used only for the purpose of administering the ACT WorkKeys assessments and is not to be copied or shared for any other purpose.

The Test Center Coordinator trains all staff members in administrative and security procedures. Each testing staff member is to be provided with a complete copy of this manual before the training session. The Test Center Coordinator works with the testing staff to ensure that the WorkKeys assessments are administered in a standardized manner in rooms that meet WorkKeys requirements.

## Section 5: Accessibility

The Workplace Documents assessment uses a variety of levels of accessibility supports including default embedded tools, open access tools, and full accommodations to allow all examinees, including those with disabilities, to participate in testing.

## 5.1 Assessment Support System

The Workplace Documents assessment contains a system of supports for effective communication that spans a continuum from the most simple, common accessibility tools used by everyone, to the most intensive accessibility supports that require the user to have specific qualifications and expertise. The levels of the continuum provide an assessment system that meets the needs of all populations tested and provides a fair communication and performance pathway for all learners.

"Accessibility is the degree to which the items or tasks on a test enable as many test takers as possible to demonstrate their standing on the target construct without being impeded by characteristics of the item that are irrelevant to the construct being measured." (AERA et al., 2014, p. 215). The Workplace Documents assessment support continuum is an inclusive concept that recognizes that the need for personalized communication supports is not restricted to any one group of examinees. It describes needs we all have, regardless of whether or not we have an official diagnostic label. It encompasses the needs of the entire testing population, including those with disabilities, those who are English Learners, as well as all the rest who have no diagnostic label at all. All of these individuals have a shared need to be able to fairly and effectively communicate what they know and can do when they take a test.

To provide a fair performance pathway for all learners, including populations with diverse needs, the development of the Workplace Documents assessment followed a theory of action known as Access by Design (Fedorchak, 2013), which incorporates elements of Universal Design for Learning (UDL)

described by the Center for Applied Special Technologies (CAST, 2011), and Evidence Centered Design (Mislevy, Almond, & Lukas, 2004; Mislevy & Haertel, 2006).

In September 2015, in anticipation of the development of this assessment, a week-long accessible test development workshop was held with leadership and content developers of WorkKeys NCRC assessments. The topic of the workshop focused on methods of mapping the characteristics and accessibility needs of learner populations to the content models intended to be measured by the NCRC assessments. During this training, accessibility consultants provided feedback with respect to accessible definitions of constructs to be tested and a plan was established for ongoing accessibility consultation and advisement during test development.

The Workplace Documents assessment accessibility supports are structured along a continuum of increasingly intensive supports designed to meet the needs of all participating learner populations. Three levels of accessibility supports are offered: (a) Embedded Tools, (b) Open Access Tools, and (c) Accommodations. Embedded tools are commonly used by many people, available to all examinees, and do not need to be requested in advance. Open Access Tools are used by fewer people, are also available to anyone, but their use must be identified and planned for locally in advance. Accommodation-level supports and tools are the most intensive levels of support. Accommodations are available to those who are qualified to use them. Currently, certain supports are only available with the paper form of the test.

Beginning in 2018, several new accessibility supports will be added to the Workplace Documents assessment for both paper and online forms. These additions will fill out the planned continuum of accessibility supports and will provide many options for unique personalization of experience for each examinee.

## 5.2 Test Administration and Accessibility Levels of Support

Over the last decade, the educational assessment profession has come to understand that all examinees have tools they need and use every day to engage in the classroom and to communicate effectively what they have learned and can do. There are different levels of support that examinees may need in order to demonstrate what they know and can do on academic tests. The Workplace Documents assessment makes several possible levels of support available. All these levels of support taken together are called accessibility supports. These accessibility supports:

- allow all examinees to gain access to effective means of communication that in turn allow them to demonstrate what they know without providing an advantage over other examinees
- enable effective and appropriate engagement, interaction, and communication of examinee knowledge and skills
- honor and measure academic content as the test developers originally intended
- remove unnecessary barriers to examinees demonstrating the content, knowledge, and skills being measured on the Workplace Documents assessment

In short, accessibility supports do nothing for the examinee academically that he or she should be doing independently; they just make interaction and communication possible and fair for each examinee.

The Workplace Documents assessment accessibility system defines four levels of support that range from minor support (default embedded system tools) to extreme support (modifications). Figure 5.1 shows the architectural structure of WorkKeys assessments accessibility supports. The width of the triangle shows the proportion of students who use that set of accessibility tools.

The Workplace Documents assessment permits the use of only those accessibility supports that validly preserve the skills and knowledge that the assessment claims to measure, while removing needless, construct-irrelevant barriers to examinee performance. The four levels of support in the assessment accessibility system represent a continuum of supports, from least intensive to most intensive, and assumes all users have communication needs that fall somewhere on this continuum. The continuum of supports permitted in the Workplace Documents assessment results in every examinee having a personalized performance opportunity.
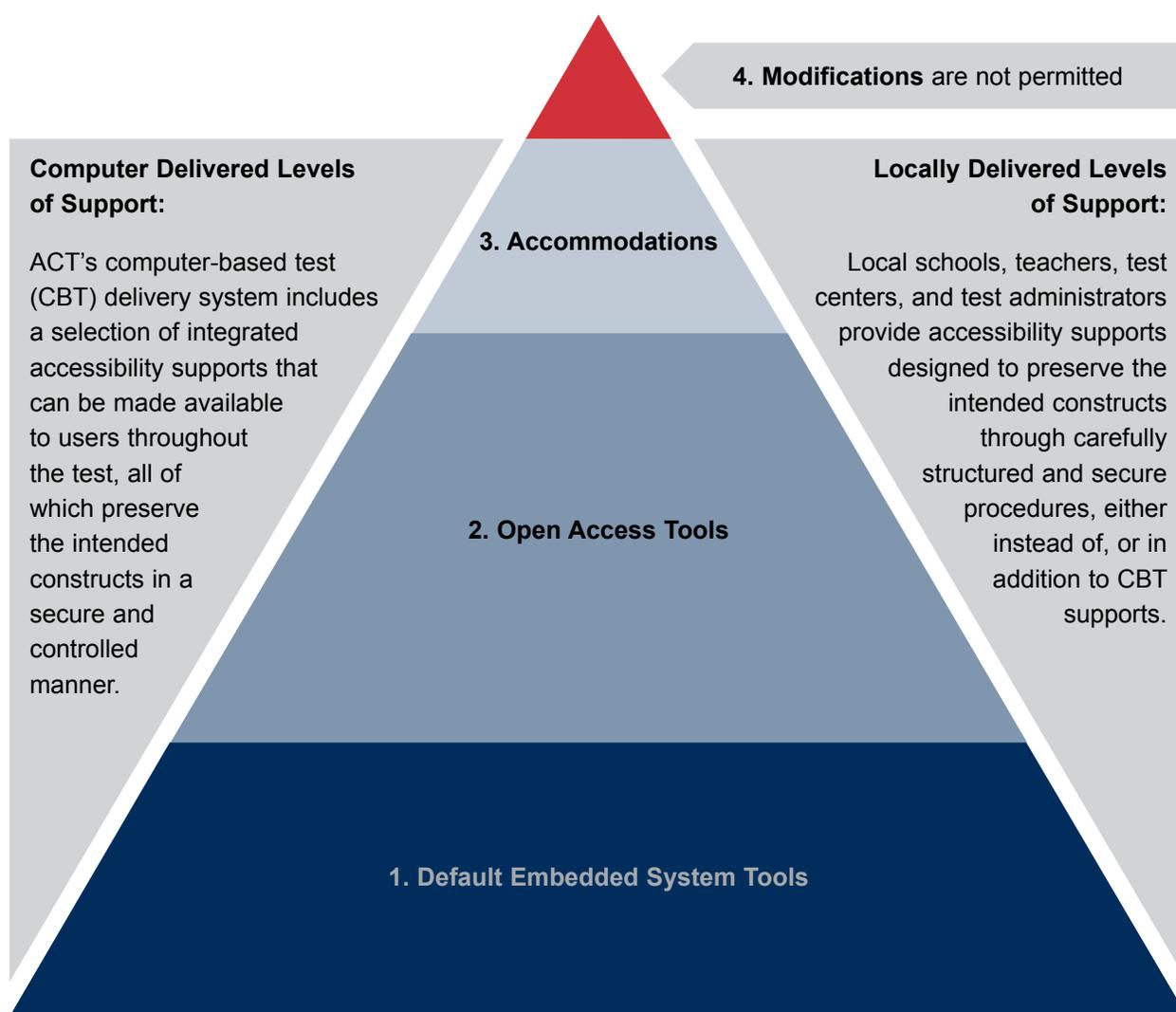
**4. Modifications** are not permitted

**Computer Delivered Levels of Support:**

ACT's computer-based test (CBT) delivery system includes a selection of integrated accessibility supports that can be made available to users throughout the test, all of which preserve the intended constructs in a secure and controlled manner.

**3. Accommodations**

**2. Open Access Tools**

**1. Default Embedded System Tools**

**Locally Delivered Levels of Support:**

Local schools, teachers, test centers, and test administrators provide accessibility supports designed to preserve the intended constructs through carefully structured and secure procedures, either instead of, or in addition to CBT supports.

**Figure 5.1: Architectural Structure of Accessibility Supports**

## Support Level 1: Default Embedded System Tools

The first level of supports is called the Default Embedded System Tools. (See the first level of the pyramid in Figure 5.1.) These tools are automatically available to a default user whose accessibility needs are sufficiently met through the basic test administration experience.

Default embedded system tools meet the common, routine accessibility needs of the most typical test takers. All examinees are provided these tools as appropriate, even examinees who have no documented support plan. Default embedded system tools include but are not limited to the following examples in online and paper tests:

- Magnifier Tool (online and paper)
- Browser Zoom Magnification (online)
- Answer Eliminator (online and paper)
- Test Directions Available on Demand (online and paper)
- Highlighter (online and paper)
- Keyboard Navigation (online)
- Mark Item for Review (online and paper)

Default embedded system tools are common supports made available to all users upon launch or start of the test; they are the accessibility tools that nearly everyone uses routinely and assumes will be made available although we seldom think of them in this way. These tools are either embedded in the basic computer test delivery platform, or locally provided as needed. No advance request is needed for these supports.

## Support Level 2: Open Access Tools

Open Access tools (see the second level of the pyramid in Figure 5.1) are available to all users, but must be identified in advance, planned for, and then selected from the menu inside the test to be activated (online), or else provided locally.

Many examinees' unique sensory and communication accessibility needs are predictable and can be met through a set of accessibility features designed into the underlying structure and delivery format of test items. Rather than overwhelm the user with all the possible tools, Open Access tools provide just the tools needed by individual users, allowing true personalization of the test experience.

Open Access tools are slightly more intensive than default embedded system tools but can be delivered in a fully standardized manner that is valid, appropriate, and personalized to the specific access needs identified for an individual examinee. Some of these require the use of tool-specific administration procedures. In the Workplace Documents assessment, Open Access tools include, but are not limited to the following examples:

- Color Contrast (online and paper)
- Line Reader (online and paper)
- Translated Verbal: Directions Only (online and paper) locally provided
- Signed Exact English (SEE) for Directions Only (paper)—locally provided

- Answer Masking (online and paper)

- Dictate Responses (online and paper)

- Respond in Test Booklet or on separate paper (online and paper)

- Audio Indicator of Time Remaining (online and paper)

- Individual Administration (online and paper)

- Special Seating/Grouping (online and paper)

Open Access tools should be chosen carefully and specifically to prevent the examinee from becoming overwhelmed or distracted during testing. Room supervisors must follow required procedures. Prior to the testing experience, examinees need to have an opportunity to practice and become familiar and comfortable using these types of tools as well as using them in combination with other tools.

## Support Level 3: Accommodations

Accommodations are high-level accessibility tools needed by relatively few examinees. (See the third level of the pyramid in Figure 5.1.) The Workplace Documents assessment system requires accommodation-level supports to be requested by education personnel on behalf of an examinee. The accommodations must be identified in advance, planned, and selected from the menu inside the test to activate them (online), or else provided locally. Accommodations often require advance ordering of specialized paper materials from ACT. The advance planning process allows any needed resources to be assigned appropriately and documented for the examinee.

Typically, examinees who receive this high level of support have a formally documented need and have therefore been identified as qualifying for resources or specialized supports that require expertise, special training, and/or extensive monitoring to select, to administer, and even to use the support effectively and securely. These can include but are not limited to the following examples:

- Braille EBAE, contracted, includes tactile graphics (paper)

- Braille UEB with Nemeth contracted, includes tactile graphics (paper)

- Cued Speech (online and paper)

- Word-to-Word Bilingual Dictionary, ACT approved (online and paper)

- English Audio DVD (designed for user with blindness) (paper)

- English Audio Reader Script (designed for user with blindness) (paper)

- Signed Exact English (SEE): Test Items

- Abacus

- Extra Time

Decisions about accommodation-level supports are typically made by an educational team on behalf of and including the examinee. Accommodation decisions are normally based on a formal, documented evaluation of specialized need and require the examinee to have personal familiarization and successful prior experience with the tools so they may be used fluidly and effectively during the test experience. Accommodation supports require substantial additional local resources or highly specialized, expert knowledge to deliver successfully and securely.

Accommodations are available to users who have been qualified by the local governing school or employment authority to use them, (e.g., a school district, a work training agency, an employer, or a branch of military or other government service). Official determination of qualification for accommodation-level support by a governing school or workforce authority is usually documented in writing in the form of an accommodation plan, or such qualification may have been routinely recognized and permitted for this examinee by that governing authority. NCRC assessments require that examinees who use accommodation-level supports have a formally documented need, as well as relevant knowledge and familiarity with these tools. Accommodations must be requested through the local test site according to the assessment procedures, as defined in the administration manual. Appropriate documentation of accommodation need, as specified in the manual, must be provided prior to testing by the examinee, or by a local governing educational authority on behalf of the examinee.

## Support Level 4: Modifications

Modifications are supports that are sometimes used during instruction, but when used in a testing situation, they alter the construct that the test is designed to measure. While they may provide an individual with the experience of taking 'a test,' modifications provide so much support that they actually prevent the examinee from having meaningful access to performance of the construct being tested. (See the top level of the pyramid in Figure 5.1.) Because modifications violate the construct being tested, they invalidate performance results and communicate low expectations of examinee achievement. Modifications are not permitted during Workplace Documents testing and if used, invalidate the resulting test score.

# 5.3 Allowable Embedded Tools, Open Access, and Accommodations

In our commitment to provide a fair testing experience for all examinees, ACT WorkKeys NCRC assessments provide an integrated system of accessibility supports that include accommodations as well as other forms (less intensive levels) of accessibility support. There are times when supports provided for those who test using the online format are combined with other types of locally provided or paper-format supports. The reverse is also true, as examinees using the paper format sometimes also take advantage of certain online options. Regardless of test format, all examinees who use accommodation-level accessibility features must have this use documented by appropriate school (or test site) personnel.

# 5.4 Valid Test Scores and Equal Benefit for All Examinees

ACT aims to ensure that all examinees may benefit equally from the Workplace Documents assessment. Accommodations and other accessibility supports administered under these standardized conditions result in a valid and fully reportable NCRC score. Use of any accessibility supports that are not specified by ACT or not properly administered violate what the test is designed to measure and result in a score that is invalid and non-comparable for the stated purposes of the assessment.

# Section 6: Test and Information Security

## 6.1 Test Security

In order to ensure the validity of the ACT WorkKeys Workplace Documents test scores, test takers, individuals that have a role in administering the tests, and those who are otherwise involved in facilitating the testing process, must strictly observe ACT's standardized testing policies, including the Test Security Principles and test security requirements. Those requirements are set forth in the *ACT WorkKeys Administration Manual—Paper Testing* and the *ACT WorkKeys Administration Manual—Online Testing* and may be supplemented by ACT from time to time with additional communications to test takers and testing staff.

ACT's test security requirements are designed to ensure that examinees have an equal opportunity to demonstrate their academic achievement and skills, that examinees who do their own work are not unfairly disadvantaged by examinees who do not, and that scores reported for each examinee are valid. Strict observation of the test security requirements is required to safeguard the validity of the results.

Testing staff must protect the confidentiality of the WorkKeys test items and responses. Testing staff should be competent and aware of their roles, including understanding ACT's test administration policies and procedures and acknowledging and avoiding conflicts of interest in their roles as test administrators for WorkKeys.

## 6.2 Information Security

ACT's Information Security program framework is based on the widely recognized ISO/IEC 27000 standard (International Organization for Standardization, 2017). This framework was selected because it covers a range of information security categories that comprehensively matches the broad perspective that ACT takes in safeguarding information assets.

ACT has developed well defined procedures and processes for the daily handling and safeguarding of secure information, as well as procedures for safeguarding secure information in the event of a disaster or adverse event. The procedures and processes are overseen by the Information Security Officer and all ACT personnel are required to participate in security training. Access to secure information by ACT personnel is limited on a "need to know basis."

ACT's Information Security Incident Response Plan (ISIRP) brings needed resources together in an organized manner to deal with an incident (classified as an adverse event) related to the safety and security of ACT networks, computer systems, and data resources.

The adverse event could come in a variety of forms: technical attacks (e.g., denial of service attack, malicious code attack, exploitation of a vulnerability), unauthorized behavior (e.g., unauthorized access to ACT systems, inappropriate usage of data, loss of physical assets containing Confidential or Confidential Restricted data), or a combination of activities. The purpose of the plan is to outline specific steps to take in the event of any information security incident.

The Information Security Incident Response Plan charters an ACT Security Incident Response Team (ISIRT) with providing an around-the-clock (i.e., 24/7) coordinated security incident response throughout ACT. Information Security management has the responsibility and authority to manage the Information Security Incident Response Team and implement necessary ISIRP actions and decisions during an incident.

# Section 7: Preliminary Field Test Data Analyses and Findings

ACT conducted a series of field test studies to better understand the psychometric properties of the Workplace Documents assessment and to make decisions related to test administration. Specifically, the studies were designed to access the following questions:

- What is the appropriate test time for the Workplace Documents assessment?
- What is the reliability (internal consistency) of the Workplace Documents assessment?
- What is the factor structure of the Workplace Documents assessment?
- Do examinee responses to the assessment items indicate that the IRT assumptions are met?
- How was the primary scale score established and what methods will be used for equating to ensure that scale scores have the same meaning regardless of form?

For each of the field test studies, ACT attempted to recruit samples representative of the regular WorkKeys testing population. Although variability existed in terms of demographic groups amongst the field test studies, each sample consisted of approximately 60% high school students and 40% adult test takers, 55% female and 45% male test takers, and 60% Caucasian, 16% African American, and 8% Hispanic test takers.

## 7.1 Testing Time Limit

ACT designed the Workplace Documents assessment to be a power test where examinees are provided adequate time to complete the assessment (Miller, Linn, & Gronlund, 2013). As such, testing time should allow sufficient time for at least 90% of examinees to complete the assessment. In the first field test study, with examinees taking assessments both on paper and online, ACT allowed examinees either 55 or 60 minutes to complete the assessment. With these time conditions, ACT surveyed examinees to learn if they believed that they had sufficient time to complete the assessment.

In the first study, ACT found that the mean testing time was a little under 30 minutes. ACT also learned that 95% of examinees completed the assessment in less than 47 minutes for both the paper and online formats. On the post-administrative surveys, 98% of the examinees who were allowed 55 minutes Agreed or Strongly Agreed with the statement—I had a sufficient amount of time to answer each test question. Based on these findings, ACT concluded that a 55-minute time limit would provide sufficient time for examinees to complete the assessment.

For the second field test study, ACT continued to evaluate the amount of time that examinees required for testing. In the second study, all examinees took two forms of the assessment online. The findings of the second field test study confirmed that 55 minutes was a reasonable time limit. In the second study, mean testing time remained approximately the same as in the first study (Form A Mean = 28.70 minutes, Form B Mean = 28.75 minutes), and 95% of examinees completed all items on the assessments in under 50 minutes. On the post-administrative surveys, slightly more than 87% of the examinees Agreed or Strongly Agreed with the statement—I had a sufficient amount of time to answer each test question.

Based on the examinee response patterns, completion rates, and survey responses, ACT concluded that allowing examinees 55 minutes to complete the assessment was reasonable and that it would result in the test not being speeded.

# 7.2 Score Reliability

ACT used Coefficient Alpha to estimate the overall reliability of the Workplace Documents assessment. Coefficient Alpha is a measure of the internal consistency of the items constituting the assessment. Prior to administration, ACT determined that 0.80 would be a minimal acceptable level of internal consistency. Following the calculation of Coefficient Alpha, ACT calculated the Standard Error of Measurement (SEM). The SEM summarizes the amount of error or inconsistency in scores on a test. Scale score SEMs were estimated using a four-parameter beta binomial model (Kolen, Hanson, & Brennan, 1992). If the distribution of measurement error is approximated by a normal distribution, true scale scores for about two-thirds of the examinees are within plus or minus one SEM from their reported scale score. Table 7.1 presents the reliability estimates and SEM based on data from the second field test study.

**Table 7.1: Workplace Documents—Reliability Estimates and Standard Error of Measurement of Form A and Form B**

| Form | N | Raw Score | | Scale Score | |
| | | Coefficient Alpha | SEM | Coefficient Alpha | SEM |
| --- | --- | --- | --- | --- | --- |
| Form A | 1,136 | 0.89 | 2.21 | 0.89 | 1.73 |
| Form B | 1,181 | 0.90 | 2.13 | 0.89 | 1.71 |

ACT also estimated the reliability of the forms applying a multivariate generalizability framework (Brennan, 2001). Generalizability theory provides estimates of universe score variance, absolute score variance, and absolute error variance. The analysis allows for the calculation of a G-coefficient, which provides an alternative estimate of internal consistency. The G-coefficients estimated for both forms of the Workplace Documents assessment was 0.90. Based on the reliability estimates and SEM, ACT concluded that the Workplace Documents met the requirements for score reliability.

# 7.3 Factor Structure

For both the second and third field test studies, ACT used Exploratory Factor Analysis (EFA) to assess dimensionality for the Workplace Documents assessment. EFA uses an inter-item correlation matrix to identify the underlying factors accounting for the observed variance in the items. In these analyses, four criteria are evaluated to determine dimensionality. A scree plot of eigenvalues can be used to determine the test dimensionality, with unidimensionality indicated by having only one eigenvalue above the "elbow" in the scree plot curve. Hatcher (1994) suggested that a factor should be retained if it accounted for at least 10% of the total variance. Reckase (1979) suggested that if the first factor explains 20% of the variance of a set of items, the item set should be considered unidimensional. Hattie (1985) maintained that the first factor is relatively strong if the factor difference ratio index (FDRI) (Johnson, Yamashiro, &

Yu, 2003) is larger than 3. FDRI is defined as the ratio of the difference between the eigenvalue of the first factor and the second factor to the difference between the eigenvalue of second and the third factor.

Table 7.2 summarizes the eigenvalues and FDRI for the Workplace Documents forms. Evaluation of the scree plots for all three forms indicated a single factor as the "elbow" appeared immediately following the first eigenvalue. Applying the rules of Hatcher (1994), Reckase (1979), and Hattie (1985) also supported the conclusion that each Workplace Documents form was unidimensional. The data presented in Table 7.2 indicate that more than 10% of the variance on all three forms could be attributed to the first factor and no other factor met the 10% threshold; that the variance attributed to the first factor exceeded 20% on all forms; and the FDRI was larger than 3 for all forms with Form C—paper being the smallest (FDRI = 9.53). Collectively, the findings provide strong evidence that the Workplace Documents assessment is a unidimensional measure.

**Table 7.2: Workplace Documents—Summary of Eigenvalues and Factor Difference Ratio Index (FDRI)**

|  | First Factor | Second Factor | Third Factor |
|---|---|---|---|
| **Form A—Online** |  |  |  |
| Eigenvalue | 12.45 (41.5%) | 1.95 (6.5%) | 1.16 (3.9%) |
| Difference |  | 10.50 | 0.79 |
| FDRI |  |  | 13.24 |
| **Form B—Online** |  |  |  |
| Eigenvalue | 12.72 (42.4%) | 1.71 (5.7%) | 1.21 (4.0%) |
| Difference |  | 11.01 | 0.50 |
| FDRI |  |  | 21.98 |
| **Form C—paper** |  |  |  |
| Eigenvalue | 12.18 (40.6%) | 2.36 (7.9%) | 1.33 (4.4%) |
| Difference |  | 9.82 | 1.03 |
| FDRI |  |  | 9.53 |
| **Form C—online** |  |  |  |
| Eigenvalue | 12.95 (43.2%) | 2.35 (7.8%) | 1.26 (4.2%) |
| Difference |  | 10.60 | 1.09 |
| FDRI |  |  | 9.72 |

*Note*. The percentage in the parenthesis is the percentage of total variance accounted for by that factor.

# 7.4 3-PL IRT Model

With evidence indicating that the Workplace Documents assessment was unidimensional, ACT checked the IRT assumption of Local Independence using the $Q_3$ method (Yen, 1984). For an item pair, $Q_3$ is the correlation of item residuals where the residual is the difference between the observed item responses

and the responses predicted for each item by an IRT model (3-PL in this study). Items that do not share a common passage are considered locally independent. The $Q_3$ indices for all items that are not in a common set were computed and were interpreted as the baseline. Following that step, $Q_3$ indices for items that share a passage were computed and compared to the baseline to evaluate whether the items in a set were more dependent than the items that were not within a set. The 95th percentile of the baseline was defined as the cut point. If the $Q_3$ for a pair of items within a set was larger than the cut point, the item pair met the assumption of Local Independence.

The $Q_3$ analysis for the Workplace Documents forms indicated that the items within a set do not show higher correlation compared to those items between sets. Evaluating the matrices led to the conclusion that no strong evidence of Local Item Dependence existed among items for all forms of the assessment.

With both IRT assumptions met, ACT analyzed item data using a 3-PL IRT model for Forms A and B for the second field test study. Table 7.3 presents the summary IRT parameter estimates for the two forms.

**Table 7.3: Summary of IRT Item Parameter Estimates**

| | IRT a | | | | IRT b | | | | IRT c | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Mean | SD | Min | Max | Mean | SD | Min | Max | Mean | SD | Min | Max |
| Form A | 1.10 | 0.38 | 0.36 | 1.86 | 0.34 | 1.01 | -1.36 | 2.29 | 0.13 | 0.06 | 0.05 | 0.29 |
| Form B | 1.09 | 0.35 | 0.45 | 2.14 | 0.27 | 1.05 | -1.51 | 3.10 | 0.13 | 0.05 | 0.06 | 0.24 |

Figure 7.1 shows the Test Characteristic Curve (TCC) and the Test Information Function (TIF) for the base form administered in the second field test study. The Workplace Documents TCC and TIF indicate an assessment that provides a good deal of measurement information for examinees who are within two standard deviations of the mean.
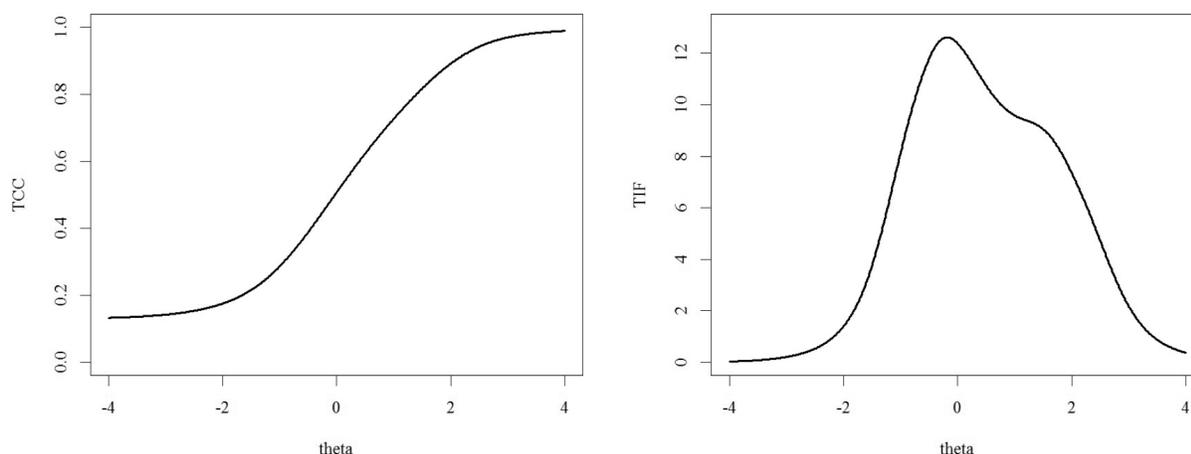


**Figure 7.1: Workplace Documents—Test Characteristic Curve and Test Information Function of Base Form**

# 7.5 Workplace Documents Score Scale and Form Equating

The base form score scale was established on the Workplace Documents assessment using examinee response data from the second field test study. The primary purpose was to develop a score scale so that examinee responses to items on all Workplace Documents forms could be converted from raw scores to scale scores, and then analyzed and compared on a common score metric. For the scaling, Form A was defined as the base form. The scaling process used IRT scaling (Ban & Lee, 2007) and applied the arcsine transformations (Kolen, 1988; Kolen & Brennan, 2004), resulting in the base form raw-to-scale score conversions. By applying the arcsine transformations, the Conditional Standard Errors of Measurement (CSEMs) were relatively equal across the score scale.

ACT set the scale scores to range from 65 to 90, which was the traditional score scale range used for the NCRC assessments. Applying the scale score average and CSEMs of the current WorkKeys 1.0 assessments, the target mean and target CSEM of the arcsine transformation were set to 77.3 and 1.7. Figure 7.2 presents the CSEMs across the full range of scale scores for the base form.
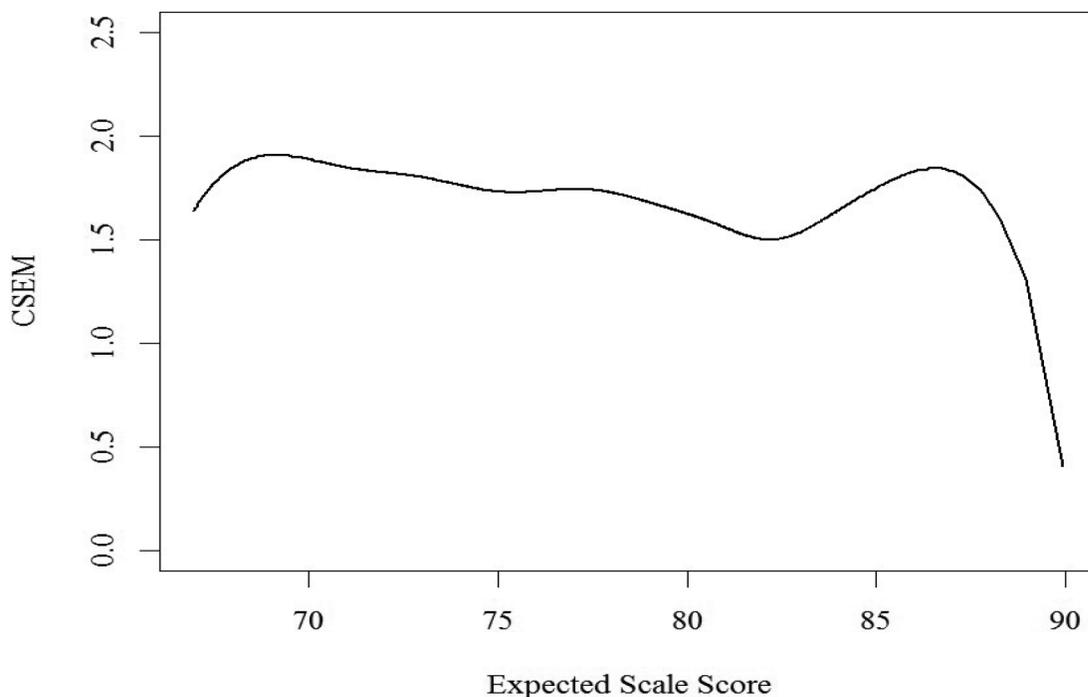


**Figure 7.2: Workplace Documents—Conditional Standard Errors of Measurement for Scale Scores**

ACT equated scores from Form B to the base form (Form A) using IRT true score equating (Kolen & Brennan, 2004). Table 7.4 provides the summary scale score statistics for Forms A and B. Based on equated scale scores, examinees who took Form B outperformed examinees who took Form A. Examinees who took Form A had slightly more variability in their scores.

**Table 7.4: Workplace Documents—Equated Scale Score Summary Statistics**

|  | N | Mean | Median | SD | Min Score | Max Score |
|---|---|---|---|---|---|---|
| Form A | 1,136 | 77.41 | 78 | 5.20 | 65 | 90 |
| Form B | 1,181 | 78.22 | 78 | 4.69 | 65 | 90 |

# Section 8: Establishing Level Score Standards

All WorkKeys assessments report to examinees their performance in terms of a scale score and a level score. The interpretation and use of the scale score is appropriate when the assessment is used to evaluate achievement or effectiveness of an educational or training program. The scale scores provide the user with finer performance distinctions than the level score. As such, they are more sensitive to gains achieved through educational or training programs.

WorkKeys level scores should be interpreted and used for selection, promotion, or other high-stakes purposes. WorkKeys assessments were originally developed to be used in conjunction with a job profiling process for employee selection and promotion. The job profiles are aligned to the assessment score levels and not to the more granular scale score.

Section 7 provided information regarding the method ACT applied to derive the Workplace Documents score scale. Following the development of the score scale, ACT conducted a standard setting in order to identify the points on the scale score where examinees who have scored at or above that point have demonstrated the ability to earn a specified level score.

## 8.1 Standard Setting Method

ACT conducted a standard setting study to establish the minimum scores required to achieve each of the five Workplace Documents levels. To establish the minimum score, ACT assembled a panel of experts consisting of educators and business people, some of whom are current WorkKeys customers. Because the Workplace Documents assessment is a criterion-referenced test, scores on the assessment are aligned to a set of skills that an examinee demonstrated. (Please refer to Section 2.4 for PLDs associated with each level score.) The goal of the standard setting process was to identify a point on the score scale where test takers who score at or above the point have demonstrated the ability to perform the skills, and test takers who score below the point have not demonstrated the ability to perform the skills.

The Mapmark with Whole Booklet Feedback standard setting method was used in the study (Schulz & Mitzel, 2005). It is a variation of the popular Bookmark procedure (Lewis, Mitzel, Mercado, & Schulz,

2012). The primary difference between Mapmark and Bookmark is the Item Map. The Ordered Item Booklet (OIB), used in both procedures, has a sample of items from the item pool ordered from easiest to hardest. The Mapmark procedure also includes an Item Map, on which the difficulty of an item is represented to an actual scale value. The Item Map, therefore, shows "how much" more difficult one item is than another item.

The data used to determine the Level Scores and Scale Scores came from the second of three field studies conducted as part of the process of updating the three assessments associated with the NCRC. Forty sites participated in the second field test study, and 2,217 test takers took one of two Workplace Documents forms. The sites included 13 high schools and 27 adult testing centers across 22 states.

Mapmark with Whole Booklet Feedback is a three-round process. In Round 1, the panelists (a) took the Workplace Documents assessment, (b) reviewed the PLDs, (c) reviewed test items and their associated Scale Score, (d) linked test items to the PLDs, and (e) placed bookmarks in the OIB for each level. Specifically, the panelists were asked to divide the items for each skill level into two groups—those items that they felt were easy enough for a minimally qualified examinee in the skill level to have mastered, and those items that were too difficult for the minimally qualified examinee to have mastered. In this context, mastery was defined as having a 2-in-3 chance of success (or a response probability of .67) on the item. This was done to establish the initial cut scores for the five levels (e.g., Levels 3–7).

For Round 2, the panelists received feedback regarding their bookmark placements relative to recommended Scale Scores on the item map and to the group's median cut score. The group was then provided with Whole Booklet Feedback. Specifically, they were provided with data showing how fifteen examinees answered each of the items on base form. Item data was provided for three examinees who scored at or near the Round 1 cut score for each skill level. The purpose was to help the panelists understand what examinees at the Round 1 cut scores "can" do and consider whether this is what examinees "should" be able to do according to each PLD. Using this feedback, panelists were asked to repeat the process of placing bookmarks in the OIB for each level.

For Round 3, panelists received feedback regarding their bookmark placement from Round 2. The feedback consisted of impact or consequential data of their Round 2 placements, which provided the percentage of examinees who achieved at or above the cut scores set for each skill level. ACT emphasized to the panelists that the PLDs should take precedence since the assessments are criterion-referenced. With that, they again placed bookmarks in the OIB.

During the final meeting, the panelists reviewed the Item Map with lines representing the Round 3 median cut scores drawn on the map. Next, they reviewed a Cut Score Distribution Chart showing the distribution of panelists' Round 3 cut scores across all skill levels. Finally, the panelists discussed impact data based on the final cut scores. The panelists approved the final median cut scores to define the five performance levels.

The Design Team reviewed the work of the Standard Setting. Design Team members evaluated whether the work of the panelists achieved the desired result of a criterion-referenced assessment with level scores aligned to the PLDs. After reviewing the panelists' work and recommendations, the Design Team approved the recommended cut scores for the five score levels of the Workplace Documents assessment.

# Section 9: Final Note Regarding On-going Studies

ACT plans to continually collect and analyze data from the Workplace Documents assessment. These analyses will range from studies of score reliability, factor structure of the assessment, fairness evaluations, and the relationship of assessment scores to significant outcome variables (i.e., job performance ratings, training completion rates and scores, program grades). ACT will periodically publish papers providing the findings of these studies. Additionally, ACT will provide information and analyses in the electronic Technical Manual in an ongoing manner as studies are completed. The information, data, and analyses are designed to provide understanding and insight regarding score interpretations and usage.

# References

ACT. (2014). *Foundational skills: What makes a skill "foundational"?* Retrieved from https://www.act.org/content/act/en/search.html?searchkey=Foundational%20Skills&submit=Search.

American Educational Research Association (AERA), American Psychological Association (APA), & National Council for Measurement in Education (NCME). (2014). *Standards for educational and psychological testing* (2014 Edition)*.* Washington: AERA Publications.

Allen, M. J., & Yen, W. M. (2002). *Introduction to measurement theory.* Long Grove, IL: Waveland Press.

Ban, J., & Lee, W. (2007). *Defining a score scale in relation to measurement error for mixed format tests.* (CASMA Research Report No. 24). Iowa City, IA: University of Iowa.

Binkley, M., Erstad, O., Herman, J., Raizen, S., Ripley, M., Miller-Ricci, M., & Rumble, M. (2012). Defining twenty-first century skills. In P. Griffin, B. McGaw, & E. Care (Eds.) *Assessment and teaching of 21st century skills* (pp. 17–66). Springer: New York.

Brennan, R. L. (2001). *Generalizability theory.* New York: Springer-Verlag.

Camara, W., O'Connor, R., Mattern, K., & Hanson, M. A. (2015). Beyond academics: A holistic framework for enhancing education and workplace success. ACT Research Report Series 2015 (4). Retrieved from http://www.act.org/content/dam/act/unsecured/documents/ACT_RR2015-4.pdf.

Center for Applied Special Technologies (CAST). (2011). Universal design for learning guidelines version 2.0. Wakefield, MA: Author. Retrieved from www.udlcenter.org/aboutudl/udlguidelines.

Crick, J. E., & Brennan, R. L. (2001). *GENOVA A general purpose analysis of variance system version 3.1*. A Fortran 77 program for analysis of variance and generalizability analyses with balanced designs.

Fedorchak, G. (2013). Access by Design—Implications for equity and excellence in education. Internal whitepaper prepared on behalf of the NH Department of Education for the Smarter Balanced Assessment Consortium. (Internal Document—no public link).

Hatcher, L. (1994). *A step-by-step approach to using the SAS system for factor analysis and structural equation modeling*. Cary, NC: SAS Institute Inc.

Hattie, J. (1985). Methodology review: Assessing unidimensionality of tests and items. *Applied Psychological Measurement*, *9*(2), 139–164.

Holland, P. W., & Wainer, H. (1993). *Differential item analysis*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Human Resources and Skills Development Canada. (2004). *Comparing classroom and workplace reading.* Ottawa: Statistics Canada and Human Resources Development Canada.

International Organization for Standardization. (2017). ISO/IEC 27000 family-information security management systems. Retrieved from https://www.iso.org/isoiec-27001-information-security.html.

Johnson, J. S., Yamashiro, A., & Yu, J. (2003). *ECPC annual report: 2002*. Ann Arbor, MI: English Language Institute, University of Michigan.

Kane, M. T. (2013). Validating the interpretations and uses of test scores. *Journal of Educational Measurement, 50*, 1–73.

Kolen, M. J. (1988). Defining score scales in relation to measurement error. *Journal of Educational Measurement, 25*, 97–110.

Kolen, M. J., & Brennan, R. L. (2004). *Test equating, scaling, and linking: Methods and practices.* New York: Springer Science and Business Media.

Kolen, M. J., Hanson, B. A., & Brennan, R. L. (1992). Conditional standard errors of measurement for scale scores. *Journal of Educational Measurement*, *29*, 285–307.

Lewis, D. M., Mitzel, H. C., Mercado, R. L., & Schulz, E. M. (2012). The bookmark standard setting procedure. In G. J. Cizek (Ed.), *Setting performance standards: Foundations, methods, and innovations* (2nd Edition) (pp. 225–253). New York: Routledge.

Lord, F. M. (1980). *Applications of item response theory to practical testing problems.* Hillsdale, NJ: Erlbaum.

Messick, S. (1989). Validity. In American Council on Education Series on Higher Education. Robert L. Linn (Ed.), *Educational measurement: 3rd edition* (pp. 13–103). New York: Macmillan.

Miller, M. D., Linn, R. L. & Gronlund, N. E. (2013). *Measurement and assessment in teaching* (11th edition). New York: Pearson.

Mislevy, R. J., Almond, R. G., & Lukas J. (2004). A brief introduction to evidence centered design (CSE Report 632). Los Angeles, CA: University of California-Los Angeles, National Center for Research on Evaluation, Standards, and Student Testing. Retrieved from files.eric.ed.gov/fulltext/ED483399.pdf.

Mislevy, R. J., & Haertel, G. (2006). Implications for evidence-centered design for educational assessment. *Educational Measurement: Issues and Practice, 25*(4), 6–20.

Mislevy, R. J., Steinberg, L. S., & Almond, R. G. (1999). *Evidence-centered assessment design.* Princeton, NJ: Educational Testing Service.

National Network of Business and Industry Associations (NNBIA). (2014). A foundation for success in the workplace: The skills all employees need, no matter where they work. Retrieved from http://businessroundtable.org/sites/default/files/Common%20Employability_asingle_fm.pdf.

Reckase, M. D. (1979). Unifactor latent trait models applied to multifactor tests: Results and implications. *Journal of Educational Statistics, 4*(3), 207–230.

Sabatini, J. (2015). *Understanding the basic reading skills of U.S. adults: Reading components in the PIAAC literacy survey.* Educational Testing Service Research Report. Princeton, NJ: ETS.

Schulz, E. M, & Mitzel, H. C. (2005, April). *The Mapmark standard setting method.* Paper presented at the annual meeting of the National Council on Measurement and Education, Montreal, Canada.

Smith, M. C., Mikulecky, L., Kibby, M. W., Dreher, M. J., & Dole, J. A. (2000). What will be the demands of literacy in the workplace in the next millennium? *Reading Research Quarterly, 35* (3), 378–383.

Yen, W. M. (1984). Effects of local item dependence on the fit and equating performance of the three parameter logistic model. *Applied Psychological Measurement, 8*, 125–145