Validity Evidence for the Cognitive ACT® WorkKeys® Assessments and the ACT® WorkKeys® National Career Readiness Certificate™

Kate E. Walton, Mary LeFebvre, and Cindy Hill

Abstract

ACT® WorkKeys® assessments measure foundational cognitive workplace skills, and the ACT® WorkKeys® National Career Readiness Certificate™ (NCRC®) is a credential test takers can earn based on their demonstration of these skills on the WorkKeys assessments. Both are used for a variety of purposes, some of which involve high-stakes decisions. This report summarizes validity evidence to date for these uses in a manner that is responsive to the *Uniform Guidelines on Employee Selection Procedures (1978)* (Equal Employment Opportunity Commission [EEOC], 1978), the *Standards for Educational and Psychological Testing* (American Educational Research Association [AERA] et al., 2014), and the *Principles for the Validation and Use of Personnel Selection Procedures* (Society for Industrial Organizational Psychology [SIOP], 2018). For organizational purposes, the report is divided into six main sections: (1) context for validation, (2) an overview of the ACT WorkKeys assessments and the ACT NCRC, (3) validity evidence based on test content, (4) validity evidence based on internal structure, (5) validity evidence based on relations to other variables, and (6) conclusion. Additional information about the assessments and NCRC can be found in the *WorkKeys NCRC Assessments Technical Manual* (ACT, 2025a).

Context for Validation

Conceptually, validity is “the degree to which evidence and theory support the interpretations of test scores for proposed uses of tests” (AERA et al., 2014, p. 11). Although the *Uniform Guidelines on Employee Selection Procedures (1978)* (EEOC, 1978), the *Standards for Educational and Psychological Testing* (AERA et al., 2014), and the *Principles for the Validation and Use of Personnel Selection Procedures* (SIOP, 2018) describe validity in slightly different terms, this report summarizes three aspects of validity outlined in the *Standards*: evidence based on test content, evidence based on internal structure, and evidence based on relations to other variables.

The cognitive WorkKeys assessments and the NCRC (which is based on the Applied Math, Graphic Literacy, and Workplace Documents assessments) are used for a variety of purposes, some of which involve high-stakes decisions. The WorkKeys assessments were primarily developed for use in employee selection and employee development. Additional uses of the individual assessments include establishing cutoffs for entry into occupational/career training

programs and measuring the effectiveness of education and workforce training programs in developing foundational work readiness skills. The primary use of the NCRC is to document an individual's level of work readiness, which is important for success in a variety of careers. Employers and industry associations use this credential as a measure of foundational skills. In this way, the NCRC serves as one of several qualifications for entry into a specific job or career path. The credential has also been used as an aggregate measure of a population's readiness for work or career (e.g., an unemployed adult population receiving career services or a census of high school students taking the assessments to meet local or state college and career readiness standards).

The *Uniform Guidelines* (EEOC, 1978) establish a common set of principles to assist employers, labor organizations, employment agencies, licensing and certification boards, and others in complying with federal law related to the use of tests and other selection procedures. Specifically, they address the use of selection procedures when adverse impact exists. While validity studies are not mandated when adverse impact is absent, such studies encourage using procedures with appropriate validation evidence to support the proposed uses of tests and other selection procedures.

ACT assessments, including WorkKeys, continue to be developed, used, and supported in a manner that strives to adhere to the professional standards and best practices reflected in the *Standards* (AERA et al., 2014) and *Principles* (SIOP, 2018). It is the responsibility of the test developer to identify the appropriate uses of its tests, as well as to warn test users against inappropriate uses when these can be anticipated. The test developer and test users also have additional responsibilities, which are addressed in the *Uniform Guidelines*, the *Standards*, and the *Principles*. These responsibilities include providing evidence and a theoretical basis to support the intended uses and interpretation of scores (AERA et al., 2014). The *Uniform Guidelines* were last revised nearly fifty years ago, and there have been significant changes in the scientific methods and professional practices related to the design, development, use, and validation of employment tests. Nevertheless, when tests are used explicitly for employment decisions, the *Uniform Guidelines* are still relevant and are in all instances the professional and technical standards to be practiced and used.

As noted in the *Standards*, not every standard applies to all tests and all uses; professional judgment is essential for determining and documenting how test developers and test users have addressed the relevant standards. A rationale and evidence may be presented when addressing standards or, similarly, when one or more standards is not applicable. Tests can often be used for additional purposes beyond the original intended use. For example, the ACT® test has been primarily used as a national college admissions test, but additional evidence has been provided to support its use in college course placement, as an indicator of college readiness, and as a means of evaluating how well schools prepare students for college (ACT, 2025c). Each additional purpose needs additional lines of evidence, which are often provided by parties other than the test developer (e.g., a test user or other researchers).

This review summarizes validation evidence in a manner responsive to the *Uniform Guidelines*, the *Standards*, and the *Principles*. The *Uniform Guidelines* define three types of validity studies

that are acceptable for providing validity evidence for a test used in a selection procedure, while the *Standards* and *Principles* consider validity a unitary concept; they focus not on the distinction between the types of validity but rather on the sources of validity evidence for a particular use. For organizational purposes, the remainder of this summary is divided into five main sections: (1) an overview of the ACT WorkKeys assessments and the ACT NCRC, (2) validity evidence based on test content, (3) validity evidence based on internal structure, (4) validity evidence based on relations to other variables, and (5) conclusion.

Overview of ACT WorkKeys Assessments

The ACT WorkKeys cognitive and noncognitive assessments currently comprise Applied Math, Graphic Literacy, Workplace Documents, Applied Technology, Business Writing, Workplace Observation, Essential Skills, and Talent. (The noncognitive WorkKeys Essential Skills and WorkKeys Talent assessments are not the focus of this report. For information on Essential Skills, see ACT, 2024. For information on Talent, see ACT, 2009). Five of the cognitive assessments are exclusively multiple-choice exams, and one (Business Writing) is a constructed response assessment. The WorkKeys Essential Skills and Talent assessments are norm-referenced, and examinees receive percentile rank scores. The cognitive WorkKeys assessments include levels, each of which was conceptualized as an independent definition of the construct to be measured, a definition not based on the psychometric properties of the assessment (McLarty & Vansickle, 1997). In this way, job analysis can be linked with the assessment of individuals so that both can address the same skill level. Although this document focuses on validity evidence for Applied Math, Graphic Literacy, and Workplace Documents, Tables 1 and 2 summarize the technical aspects of each assessment.

ACT WorkKeys assessments can help employers identify a pool of qualified applicants who have achieved the levels of proficiency needed to perform a job as determined through job analysis. The assessments should be used in combination with additional measures (e.g., tests, interviews, or other selection procedures) that the employer deems appropriate and relevant for employee selection or other employment decisions.

Table 1. Summary of ACT WorkKeys Cognitive Assessments' Characteristics

Test name	Construct	Content	Item type	Total number of items	Score type	Timing	Format	Language
Applied Math	Applying mathematical reasoning, critical thinking, and problem-solving techniques to work-related problems	Examinees set up and solve problems and do the types of calculations that occur in the workplace.	Multiple-choice	34	Levels 3–7 Scale score (65–90)	55 min. English 70 min. Spanish	Paper and online	English and Spanish
Graphic Literacy	Working with workplace graphics	Examinees find information in a graphic and compare, summarize, and analyze information found in related graphics.	Multiple-choice	38	Levels 3–7 Scale score (65–90)	55 min. English 70 min. Spanish	Paper and online	English and Spanish
Workplace Documents	Reading and using written texts to do a job	The written texts include memos, letters, directions, signs, notices, bulletins, policies, and regulations.	Multiple-choice	35	Levels 3–7 Scale score (65–90)	55 min. English 70 min. Spanish	Paper and online	English and Spanish
Business Writing	Writing an original response to a work-related situation	Components include sentence structure, mechanics, grammar, word usage, tone and word choice, organization and focus, and development of ideas.	Constructed response	1 prompt	Levels 1–5 Scale score (50–90)	30 min.	Paper and online	English

Test name	Construct	Content	Item type	Total number of items	Score type	Timing	Format	Language
Workplace Observation	Observing, following, understanding, and evaluating processes, demonstrations, and other on-the-job procedures	Examinees must focus and notice what they are observing while also following, interpreting, synthesizing, analyzing, or evaluating what they have observed.	Multiple-choice	35	Levels 1–5	55 min.	Online	English
Applied Technology	Solving problems with machines and equipment found in the workplace	Examinees analyze and solve a problem that may involve the application of four areas of technology: electricity, mechanics, fluid dynamics, and thermodynamics. They apply existing tools, materials, or methods to new situations.	Multiple-choice	34	Levels 3–6 Scale score (65–90)	55 min.	Paper and online	English and Spanish

Table 2. Summary of ACT WorkKeys Noncognitive Assessments' Characteristics

Test name	Construct	Content	Item type	Total number of items	Score type	Timing	Format	Language
Essential Skills	Work ethic, collaboration, resilience, creativity, leadership, integrity	The test assesses individuals' standing on six essential skills.	Likert items and situational judgment tests (SJTs)	36 Likert items and 12 SJTs	Percentiles	20 min.	Online	English
Talent	Carefulness, cooperation, creativity, discipline, goodwill, influence, optimism, order, savvy, sociability, stability, striving	The assessment measures personality characteristics in the workplace.	Likert items	165	Percentiles	35 min.	Online	English

ACT NCRC

The ACT NCRC is a portable credential that demonstrates test takers' foundational work skills related to Applied Math, Graphic Literacy, and Workplace Documents. Foundational skills are defined as the fundamental portable skills that are critical to training and workplace success (Clark, 2015). The NCRC is issued at four levels based on examinee level scores on the component tests (see Table 3).

Table 3. NCRC Levels

Certificate level	Level score
Platinum	6 or above on all three tests
Gold	5 or above on all three tests
Silver	4 or above on all three tests
Bronze	3 or above on all three tests

In 2003, several states originated the NCRC's predecessor, the Career Readiness Certificate (CRC; Bolin, 2005). States could independently decide whether to include the WorkKeys assessments (then called Applied Mathematics, Locating Information, and Reading for Information) for the state-issued CRCs; they could also define the certificate level as being equivalent to the lowest score obtained across the three assessments. In 2007, ACT began issuing a national version of the certificate using the same WorkKeys assessments and levels.

In response to evolving workplace demands and technological advancements in the 21st century, a comprehensive review of the WorkKeys assessments was launched in 2014, leading to updates in 2017 that better reflect how foundational skills are applied on the job today. Applied Mathematics was renamed Applied Math and includes scenarios such as identifying the correct formula to use in a spreadsheet. Locating Information was retired and replaced by Graphic Literacy, which emphasizes interpreting modern workplace visuals—like dashboards that summarize complex data—and selecting the most effective graphic for a specific purpose. Reading for Information became Workplace Documents, now including tasks like reading a string of emails to determine the appropriate action. For clarity and consistency, the assessments will be referred to by their current names—Applied Math, Graphic Literacy, and Workplace Documents—throughout the remainder of this report.

The NCRC is intended to be used in a variety of ways:

- Employers may recommend that applicants provide their NCRC level in addition to traditional criteria (e.g., employment application, credentials, or interview) as part of the job application process. In these instances, employers may then use the NCRC and the other criteria to screen prospective applicants for hiring. In this scenario, the employer is using the NCRC and other criteria to identify a qualified pool of applicants[1] and is not requiring a specific NCRC level.

- Employers may use the NCRC to make employment decisions and require a specific level (e.g., Gold). In such instances, a formal job profile should be conducted. When

available, evidence may be transported from job analysis studies of similar positions requiring the same skills. In this situation, the NCRC level is used to partially determine whether an applicant is qualified, and additional evidence that the specified level represents skills required on the job should be documented.

- States, communities, and schools may use the NCRC to document an individual's level of key work readiness skills. Specifically, state and local workforce and education agencies often provide NCRC testing so individuals can document their readiness for a potential job or career path. In this scenario, states, communities, or institutions use NCRC testing to help individuals improve their work readiness skills.

- Finally, states, communities, or schools may use the NCRC to document the aggregate work readiness of a community, region, or state. As part of a Work Ready Community™ initiative, education, economic, and workforce partners provide NCRC testing to gauge the skill levels of different segments of the local labor market. Used in this manner, the NCRC can help workforce and education stakeholders identify skill gaps in both the emerging and current labor forces at the regional or state level.

Overall Claims and Interpretive Argument of ACT WorkKeys Test Scores

ACT WorkKeys assessments can be used for (1) pre-employment screening to identify individuals who have achieved the necessary levels of skill proficiency, (2) pre-employment screening to identify less desirable candidates based on behaviors associated with poor job performance, (3) employee development, and (4) determining the appropriate level of fit with occupations in terms of interests.

When WorkKeys tests are used for pre-employment screening or other high-stakes employment decisions, employers should demonstrate that the knowledge and skills in the pre-employment measure are linked to work behaviors and job tasks either through job profiling or through research that links the test to job performance. In addition, employers should require other traditional criteria, such as job applications, credentials, interviews, and references, when evaluating job applicants. When WorkKeys tests are used for employee development or to assess the readiness of either individuals or groups, criteria other than job performance may be more relevant (e.g., individual earnings, employment, or training completion). If cutoff scores are needed, they should be established at the appropriate levels, and the process for determining those levels should be clearly documented (AERA et al., 2014; SIOP, 2018).

The NCRC can be used by (1) an employer to make employment decisions based on a specific level (e.g., Gold), (2) states, communities, and schools to document an individual's level of foundational career readiness, and (3) states, communities, or schools to document the aggregate work readiness of a community, region, or state.

Evidence Based on Test Content

An analysis of the links between test content and the construct it is intended to measure, such as reading ability or conscientiousness, can provide important validity evidence. The process of compiling construct evidence starts with skill identification and test development, in that items are written to clearly map to the target constructs being measured. It continues with research conducted to assess alignment, which entails evaluating the correspondence between test content and student learning standards (AERA et al., 2014). Here we discuss the identification of WorkKeys skills, the test development process, job profiling, and alignment studies.

Skill Identification

In developing the WorkKeys cognitive assessments, ACT consulted workforce developers, employers, and educators to identify foundational workplace skills that (1) are used in a wide range of jobs, (2) could be taught in a short period of time, and (3) could be determined through job analysis (ACT, 1992). Initial WorkKeys skills were selected using data about high-demand skills identified by employers (ACT, 1987; Agency for Instructional Technology, 1989; Bailey, 1990; Carnevale et al., 1990; Center for Occupational Research and Development, 1990; Conover Company, 1991; Electronic Selection Systems Corporation, 1992; Greenan, 1983; Kane et al., 1990; Tatzel & Lamdin, 1975). ACT gathered survey data from employers and educators in seven states (Illinois, Iowa, Michigan, Ohio, Oregon, Tennessee, and Wisconsin) and from several community colleges in California. Collectively, representatives from each state and organization served as charter members in defining development efforts for WorkKeys.[2] Charter members also assisted ACT in both the design and review of plans and materials, as well as in providing examinee samples to test prototypes and take field-tests.

Based on extensive reviews of the literature and empirical data collected from hundreds of educators and employers, 11 skills were initially selected for the WorkKeys battery of assessments (McLarty, 1992). As with all assessments and the content and skills they measure, the WorkKeys assessments need to be continually reviewed to confirm that the content assessed remains relevant and aligned to foundational workplace skills, which are evolving over time, and that the constructs measured represent key skills and behaviors required across jobs.

Test Development

Understanding the relationship between WorkKeys level scores and skill definitions is critical for both understanding WorkKeys test development and interpreting WorkKeys scores. ACT subject matter experts (SMEs) established the skill levels for the cognitive skills before the assessments and scale scores were developed. The level scores are interpreted similarly to performance levels or achievement levels reported on educational tests. Further, the skill levels are used both to place examinees into performance categories and to identify the skill requirements for specific jobs (McLarty & Vansickle, 1997). ACT designed the level scores to be easily interpretable so as to facilitate proper use in the selection, promotion, development, and classification of individuals or groups. The WorkKeys assessments also report a more granular

scale score. Scale scores provide finer distinctions between examinees and are useful for program evaluation, group comparisons, and research studies.

Panels of employers and educators initially developed the WorkKeys skill levels. They began by defining the level of performance required for each skill based on available data and expert judgment. Next, they identified a list of example tasks within each skill domain and selected those tasks most critical for performance across a variety of jobs. The SMEs then ordered the tasks by difficulty and complexity, building a hierarchy of skills with the easiest and least complex tasks forming the lowest skill levels and the harder and more complex tasks forming the highest skill levels. Using information provided by the skill hierarchies, ACT developed initial test blueprints (McLarty & Vansickle, 1997). ACT developed the WorkKeys cognitive skill levels to confirm that (1) each skill was assessed in a manner consistent with how it was used in the workplace, (2) the lowest level of a skill was set at approximately the lowest level at which an employer would wish to set a standard, (3) the highest level of a skill was set at the highest level that entry-level jobs would require, and (4) the steps between levels were large enough to be distinguishable and small enough to be meaningful (McLarty, 1992).

Skill levels are also critical for item development. Based on the defined skill levels, ACT first developed item prototypes and then used the prototypes as models for item development. Item writers receive training and coaching from ACT content specialists and develop items aligned to specific skill levels (e.g., a writer may develop an item to measure Reading for Information skills at Level 4). ACT provides item writers with guides that carefully define the skill levels and state that each item must align to an identified skill.

When writers complete and submit items, ACT also requires them to submit workplace justifications demonstrating that each item assesses the skill in a manner consistent with the way that skill is used in the workplace. Items developed and selected for use on the WorkKeys assessments go through several reviews that confirm that items are job-related and fair to candidates. External item reviewers carefully scrutinize each item for job relatedness and examinee fairness. They determine whether the item is acceptable based on the following:

- Does the item assess skills needed in the workforce?
- Is the item applicable across a wide range of jobs and occupations?
- Does the item give an unfair advantage to examinees in some occupations or penalize examinees in other occupations?
- Does the item contain any content that might be offensive from a cultural, racial, or gender perspective?
- Are the knowledge, skills, and information required to answer the item equally available to all demographic groups?

Test blueprints articulate the complexity level required for items at each level. All WorkKeys foundational assessments incorporate hierarchies of increasingly complex tasks and skills. Using the skill hierarchies, ACT originally based the scoring rules on a theoretical Guttmann model (Guttmann, 1950; McLarty, 1992). Because of the limitations of Guttmann scaling and the resulting complexities in applying it to actual test data, in 1996, ACT transitioned from the Guttmann item-based pattern scoring model to a probabilistic model based on total score (the item response theory [IRT] model).

ACT was able to transition to an IRT scoring model because of two important characteristics of the WorkKeys assessments. First, test items are based on a hierarchy of difficulty and complexity. Second, the assessments are unidimensional; that is, the test blueprints and items each measure a single domain. For these two reasons, all the test items (not just the items representing a specific mastery level) can be used to infer an examinee's mastery of the skills at a defined level. To illustrate, an examinee's responses to all 35 items on the Workplace Documents assessment can be used to determine whether he or she has mastered the skills at Level 3. IRT methods are used to analyze item responses and provide an estimate of the examinee's ability that does not depend on a specific test form or set of items. Based on the IRT ability estimate, it is possible to estimate an examinee's probability of correctly answering any item in the pool. After estimating an examinee's ability on the IRT scale, ACT can estimate the examinee's expected performance (percentage of items answered correctly) with respect to the entire pool of items at each level. ACT used this information to establish cutoff scores for each level; these scores represented an examinee being able to answer 80% or more of the items correctly within a given level (Schulz et al., 1997).

The advantage of using the IRT total score model over the Guttmann item-based pattern scoring model is that it provides higher levels of reliability. Because the Guttmann model was based on only the items administered at a given level (e.g., six, eight, or nine items), score reliability was relatively low. The IRT model uses all the items administered on the form (e.g., 30 or 32 items) to estimate the examinee's mastery. Using more information to determine test scores normally provides better score reliability (Schulz et al., 1997).

ACT has developed equivalent forms of its cognitive assessments to maintain test security. Each new form is constructed to adhere to the test blueprint. All WorkKeys test blueprints define both content requirements and statistical requirements. To control for the inevitable small differences in form difficulty and to confirm the accuracy and meaningfulness of scores, ACT equates and scales each newly developed form to a base form. Test equating is a statistical process used to verify that scores obtained from a new form have the same meaning and interpretation as scores earned on earlier forms (Kolen & Brennan, 2004). Equating indicates whether examinees' scores are independent of the test form. It should not matter to examinees which form they take (Lord, 1980); more specifically, they should have the same probability of earning a specific score regardless of whether they take Form X or Form Y.

To enhance score interpretations and decrease potential score misuse, standardized assessments develop a score scale that is independent of the number or percentage of items answered correctly (i.e., the raw score). The score scale provides a common metric through

which performance on different forms of the assessment can be compared. By establishing a common metric and then applying IRT equating methods to place raw scores on the common scale, test developers give scores earned on different forms the same meaning.

ACT WorkKeys multiple-choice cognitive assessments have a 26-point score scale, which ranges from 65 to 90. The assessments provide level scores with either four or five levels. Through the equating and scaling process, test users are able to interpret and compare scores achieved on two different forms of the same assessment (e.g., Applied Math or Workplace Documents).

In addition to equating and scaling, ACT continually evaluates the reliability of WorkKeys assessment scores using a variety of techniques, including estimating the internal consistency of each test form, conducting generalizability analyses, computing scale score reliability estimates, and estimating classification consistency. Classification consistency refers to the extent to which classifications of examinees agree when obtained from two independent administrations of a test or two alternate forms of a test.

Testing accommodations are available for individuals with disabilities taking the WorkKeys tests, as required by the Americans With Disabilities Act. Accommodations are authorized by the test supervisor, following ACT guidelines and with proper documentation, and may include the use of special testing materials provided by ACT, such as large-print test booklets, large-print answer documents, braille versions of the tests, and reader's scripts. ACT (2025b) provides an accessibility support guide.

Job Profiling

Overview

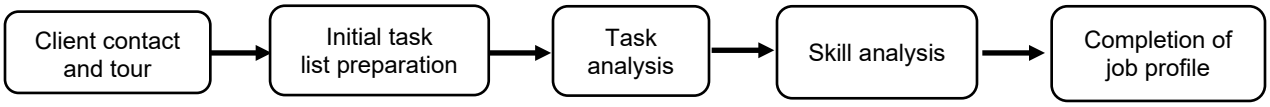
The *Uniform Guidelines* (EEOC, 1978), the *Standards* (AERA et al., 2014), and the *Principles* (SIOP, 2018) all indicate that knowledge and skills in pre-employment measures should be demonstrably linked to work behaviors and job tasks. Both the *Standards* and *Principles* suggest that expert judgment can be used to determine the importance of job tasks and to relate such tasks to the content domain of a measure. This process is commonly conducted through a job analysis, which identifies the tasks required to do a job; a job analysis also helps ensure content validity during the development of the content blueprint and test items (Cascio, 1982; Dunnette & Hough, 1990). The WorkKeys assessments were designed to assess generalizable skills and skill levels associated with many jobs. As such, the content-related validity evidence for the WorkKeys assessments was originally established by SMEs across numerous jobs who linked WorkKeys skills and skill levels to specific tasks and job behaviors for a particular job.

ACT employs a job profiling procedure that focuses on the skills and behaviors present across the WorkKeys assessments. It is a multistep process that includes the creation of one or more groups of SMEs, who are typically job incumbents. Ideally, the SMEs make up a representative sample across a variety of demographic factors (e.g., race, ethnicity, gender, geographic location). Incumbent SMEs with disabilities also participate in the focus groups (in such cases, profilers are encouraged to contact ACT for recommendations for accommodations during the

profiling process). A job profiler who has been trained and authorized by ACT conducts the process and completes the content validity report. Each profile represents a content validation study at the organizational level.

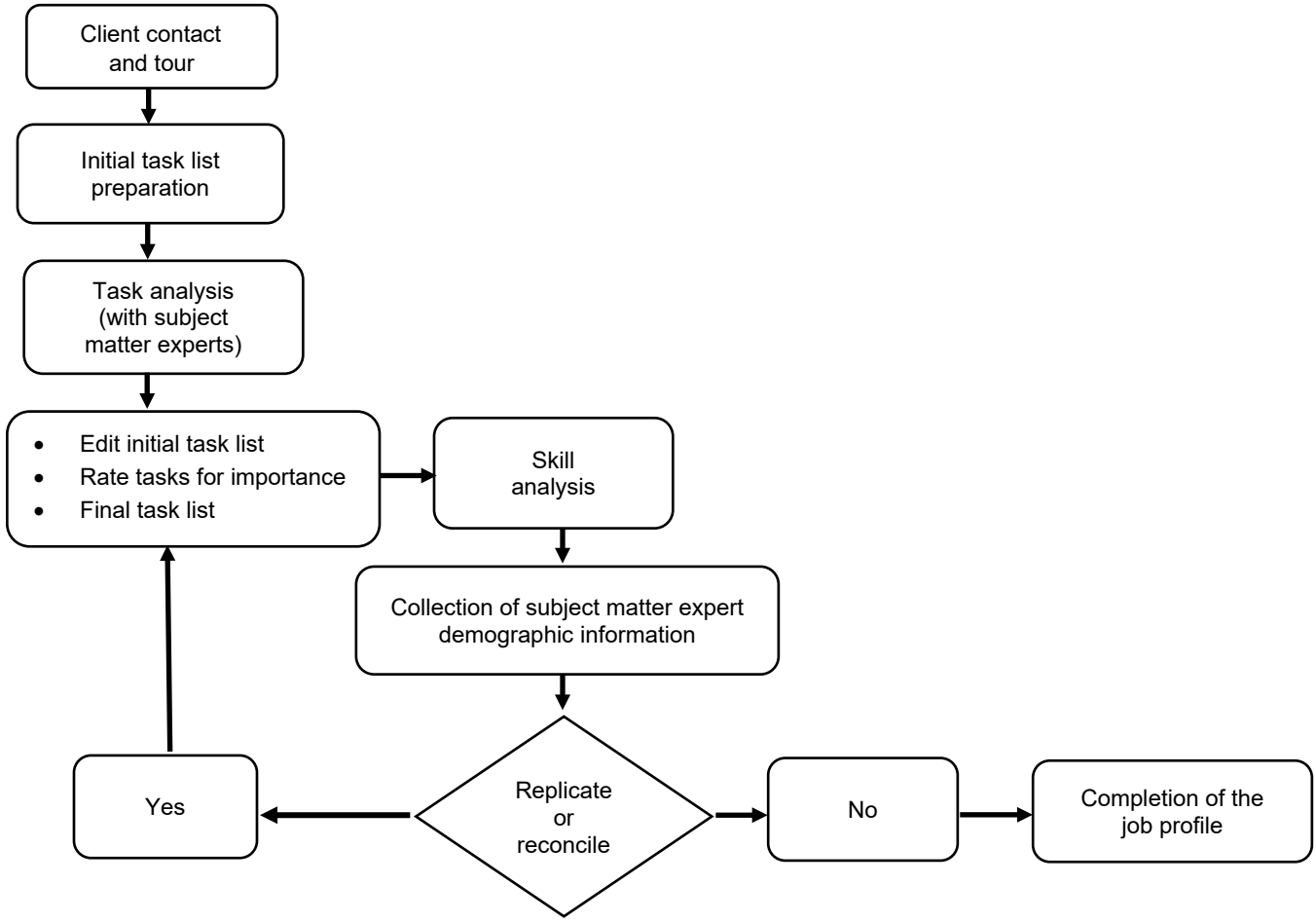
Figure 1 provides an overview of the job profiling process. Note that the task and skill analyses can be replicated by more than one group.

Figure 1. Job Profiling Process



The process begins with an ACT-authorized job profiler consulting with a client to understand their needs. This includes conducting a site tour, shadowing employees on the job, interviewing employees, and researching the job to develop a preliminary task list. The profiler then facilitates a task analysis session with a group of SMEs, who review and refine the task list to ensure it accurately reflects the responsibilities attached to the position within the organization. Following this, the SMEs rate each task based on its importance. Only those tasks identified as important by a majority of the SMEs are included in the discussions during the next step, skill analysis. Figure 2 identifies the outcomes of the task analysis and their place within the complete job profiling process.

Figure 2. Task Analysis Process



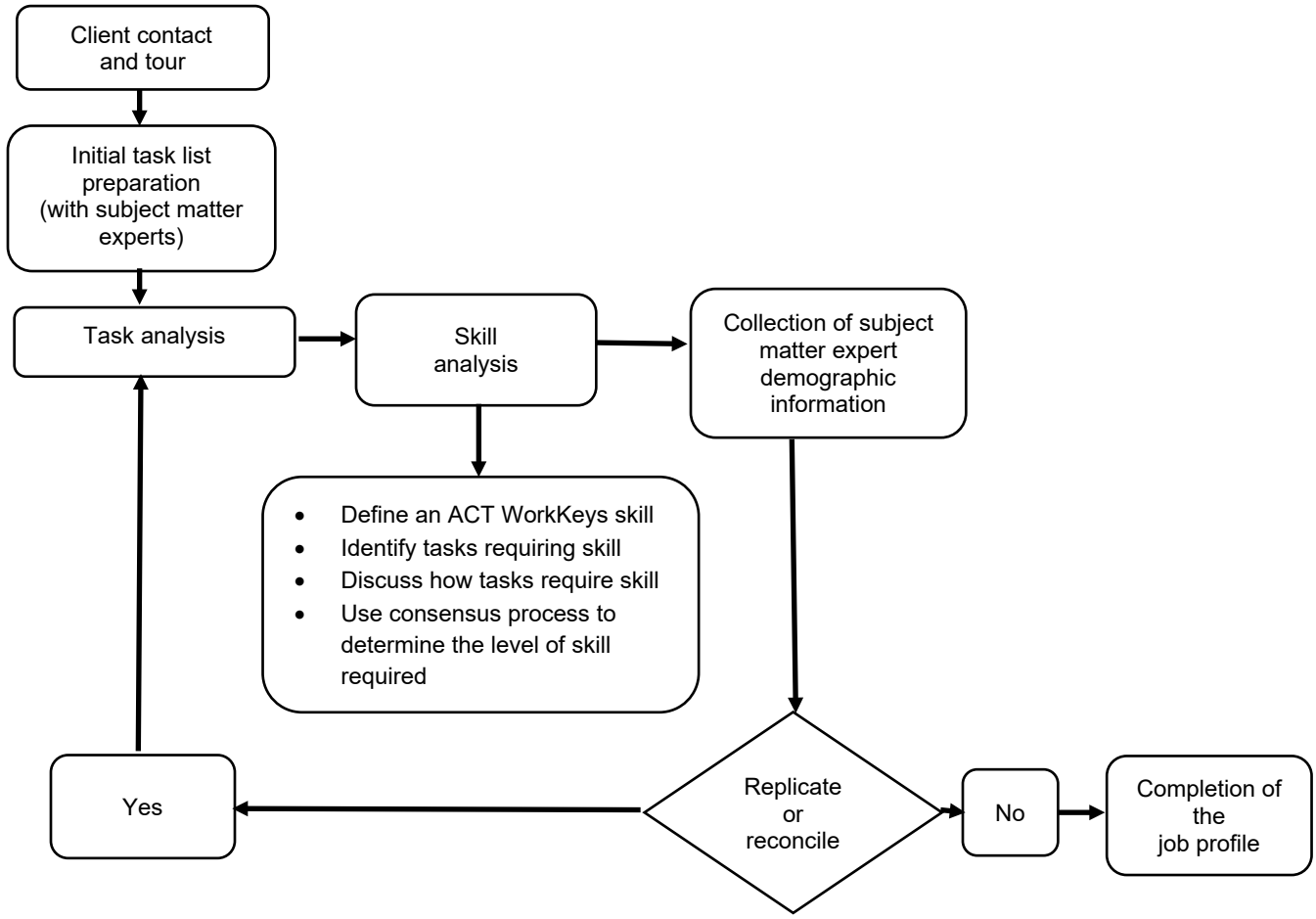
Equally important is the skill analysis, where the SMEs review each skill measured by the WorkKeys assessments. Once the SMEs understand the definition of a WorkKeys skill and have determined its relevance to the job, they independently identify the important tasks on the final task list that require the skill. They also identify how the skill is specifically used to complete each task. After discussing the relationship between the skill and the important tasks, the SMEs reach a consensus about the level of skill required for the job.

As part of the skill analysis, SMEs use a structured, iterative, level-by-level process to determine the skill level required for the final set of tasks. WorkKeys skill levels are ordered by increasing difficulty. SMEs usually begin at the lowest skill level (unless the job clearly requires higher-level skills) and evaluate whether job requirements fall below, at, or above that level, moving upward or downward across adjacent levels as needed.

A final decision is made only after SMEs have considered at least one level above and one level below the level set (unless the judgment occurs at the highest or lowest level available). In some cases, the required skill level falls outside the range measured by WorkKeys.

Figure 3 identifies the outcomes of the skill analysis and their place within the complete job profiling process.

Figure 3. Skill Analysis Process



Summary of Evidence

Since its inception, the WorkKeys job profiling process has been widely adopted across industries. To date, over 6,000 organizations have conducted more than 21,000 job profiles across 48 states, in the District of Columbia, and internationally in countries such as Canada, Germany, Finland, and Sweden—often in support of companies establishing new operations in the United States. This means that more than 88,000 employees have directly linked their job tasks to the foundational skills measured by WorkKeys. These data reinforce ACT's position that WorkKeys skills are foundational, fundamental, and portable. That is, they are key for conveying and receiving information critical to workplace training and success, they serve as a base for learning and performing more complex tasks, and they are transferable across a wide range of occupations rather than being job-specific. Foundational skills are also necessary for acquiring more specialized knowledge and skills, a pattern observed across both skilled trades and professional roles.

As of June 2025, a total of 21,811 ACT job profiles had been completed, adding to the content validity evidence base for the WorkKeys assessments. ACT's job profile database, JobPro, reflects a broad cross section of the labor market, covering 57% (568) of the 998 O*NET codes (excluding military occupations). It also includes foundational skills data for 195 (53%) of the 369 Bright Outlook Occupations, as defined by ONET using U.S. Bureau of Labor Statistics projections for 2024. JobPro represents occupations across all major occupational families and spans a range of educational and training levels required for job entry. Figure 4 illustrates the distribution of a cohort of job profiles by major occupational family, while Figure 5 shows the percentage of occupational codes with job profiles within each 2024 career cluster.

Figure 4. Distribution of Job Profiles and U.S. Employment by Occupational Family

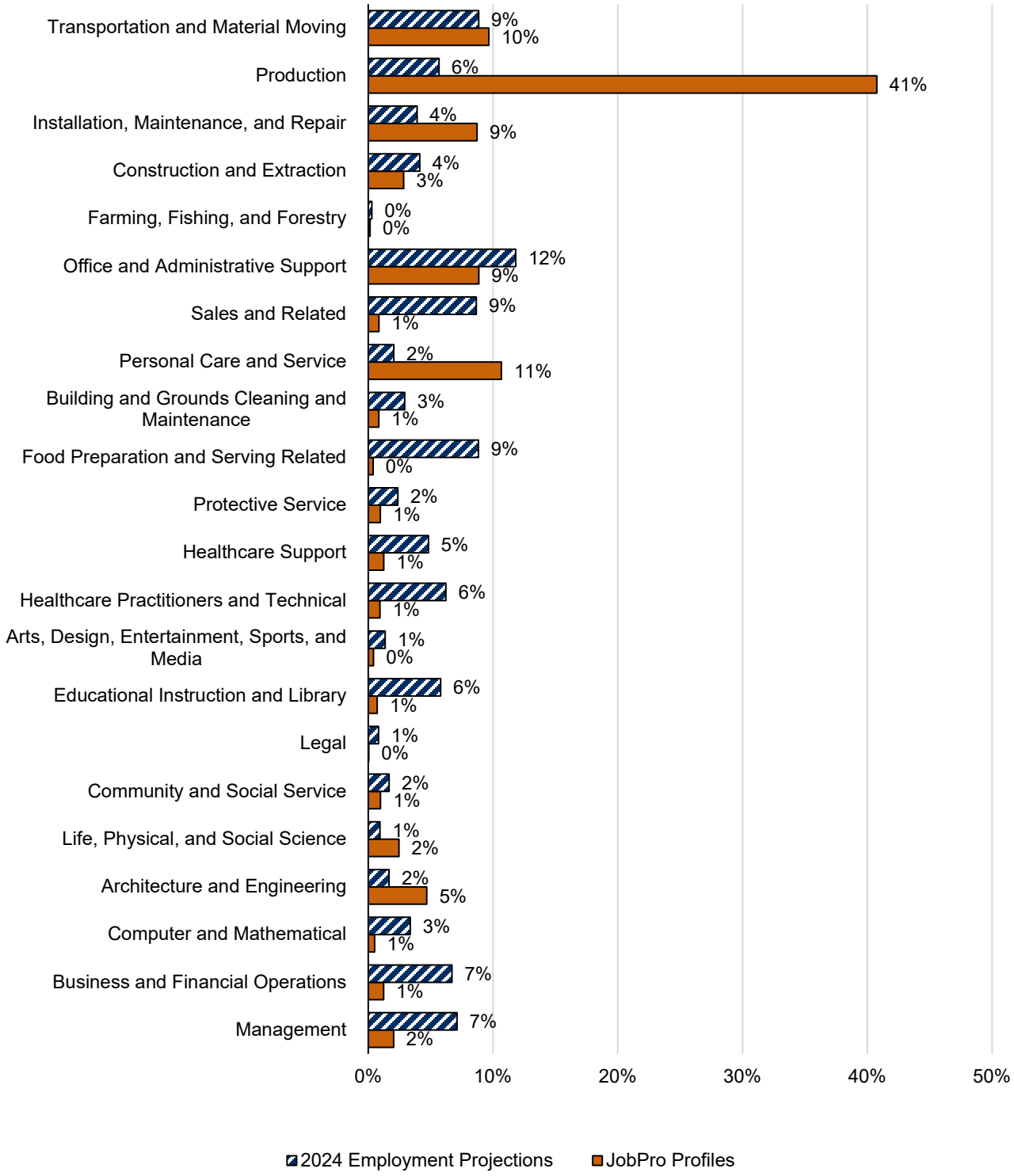
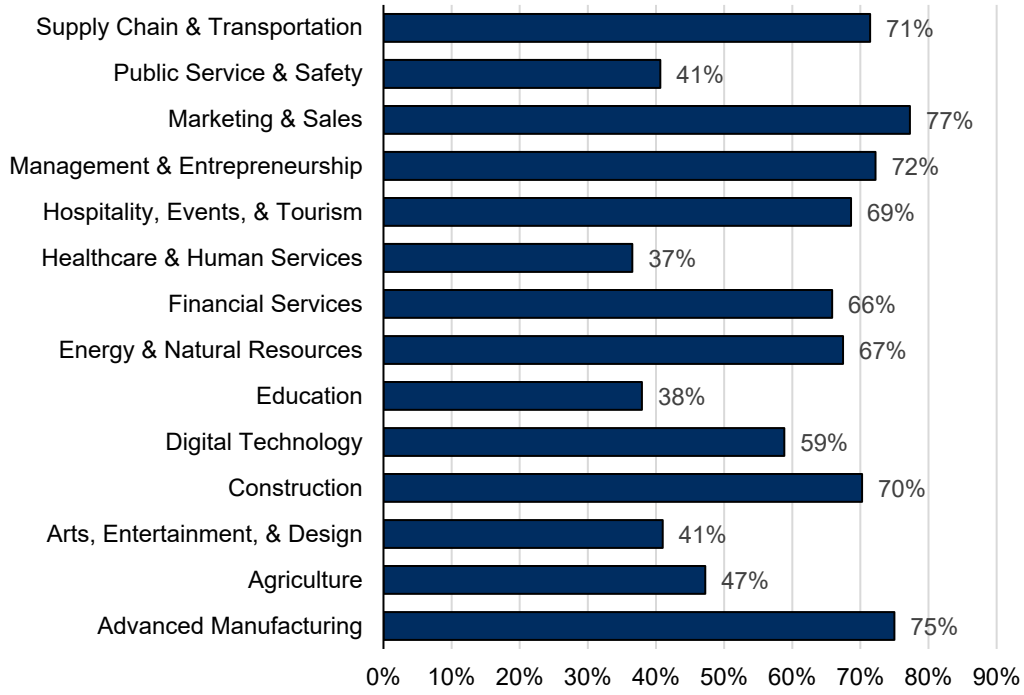


Figure 5. Percentage of O*NET Occupation Codes With WorkKeys Job Profiles in Each Career Cluster



Alignment Studies

Alignment studies, or crosswalks, provide additional test content validity evidence by evaluating the correspondence between a skill or learning standard and the content outline and its items. Various alignment studies of the WorkKeys assessments have been conducted over the years to compare the skills included in secondary education standards and business competency models.

In 2014, ACT conducted an alignment study comparing the WorkKeys assessments with a preliminary version of an industry competency model developed by the National Network of Business and Industry Association (NNBIA; ACT, 2014a). The competency model, Blueprint, was developed to describe the common employability skills necessary for jobs across all industries and sectors. The NNBIA Blueprint includes four competency buckets—Personal Skills, People Skills, Applied Knowledge, and Workplace Skills—with core skills highlighted within each bucket. Eleven WorkKeys cognitive and noncognitive assessments were included in the alignment study. The analysis found a high level of correspondence between the NNBIA Blueprint and the WorkKeys assessments. Roughly 90% (70 out of 79) of the common employability skills identified in Blueprint were measured by WorkKeys. The greatest deficiency areas were science and technology knowledge and business fundamentals.

ACT (Hill & Lyons, 2017) also explored the connection between WorkKeys competencies and the beta Connecting Credentials Framework, which has eight levels indicating the relative complexity, breadth, and depth of a competency. It was designed to connect competencies from diverse credentials, instructional courses, etc. by using a common proficiency language to describe what test takers at each level should know and what skills they should have. Notably, the study found a tremendous amount of congruence between the NCRC and the Framework in terms of the progression of levels for various competency statements (e.g., the Applied Math skill of converting from a decimal to a common fraction is a Level 3 in WorkKeys and a Level 1 in the Framework). The crosswalk showed that for Applied Math, Graphic Literacy, and Workplace Documents, there is a direct relationship between the proficiency levels of the Framework and those of the WorkKeys leveling model.

In 2021, ACT conducted a crosswalk between the three NCRC assessments and the College and Career Readiness Standards for Adult Education (CCRSAE), a subset of the Common Core State Standards considered most relevant to adult education (ACT, 2021). The crosswalk shows that the WorkKeys NCRC suite of assessments is designed to measure skills and knowledge that directly contribute to employability and work success, which are integral to the CCRSAE. Moreover, given that the National Reporting System's Educational Functioning Levels for Language Arts and Mathematics also use the CCRSAE as their foundation, the results can be extended to demonstrate a crosswalk to the National Reporting System's Educational Functioning Levels. As an example, everything assessed by the Applied Math assessment is covered by the CCRSAE standards, and the majority of the CCRSAE Mathematics Level A and B standards crosswalk to Applied Math Levels 3 and 4, while the majority of the Level D and E standards crosswalk to Applied Math Levels 5–7.

Separate alignment studies were conducted between the WorkKeys assessments and the 8th- and 12th-grade National Assessment of Educational Progress (NAEP), a national assessment of what students know and can do in various academic subjects. Both studies examined the alignment of WorkKeys Reading for Information and Applied Mathematics and were conducted to determine whether NAEP could support inferences about career readiness. WorkKeys was selected by the National Assessment Governing Board (NAGB) because it was considered a national indicator of career readiness (NAGB, 2010a). "While NAEP has been designed to provide evidence of what students in the United States know and can do in a broad academic sense, WorkKeys assessments provide information about job-related skills that can be used in the selection, hiring, training, and development of employees" (Dickinson et al., 2014, p. i). The alignment between WorkKeys Applied Mathematics and the NAEP Grade 12 Mathematics assessment was conducted in 2010 by the NAGB (NAGB, 2010a). The study investigated the alignment between the test items and levels in WorkKeys Applied Mathematics and the items and five content strands[3] in the NAEP Grade 12 Mathematics framework. Alignment was found for 75% of the NAEP items, while 40% of the WorkKeys Applied Mathematics items aligned to the NAEP standards. The NAEP Mathematics items with the strongest alignment to Applied Mathematics included problem-solving applications of number operations and measurement. NAEP Mathematics items to which no Applied Mathematics items aligned were related to geometry, data analysis, statistics, probability, and algebra. The alignment study of WorkKeys Reading for Information and NAEP Reading found some degree of alignment with NAEP

standards on locating/recalling and integrating/interpreting (NAGB, 2010b). Reading for Information items to which no NAEP Reading items aligned included items about applying complex multistep conditional instructions to similar and new workplace situations; determining the meaning of work-related acronyms, jargon, and technical terms; and figuring out general principles contained in informational documents and applying them to similar and new workplace situations.

In 2014, the NAGB conducted another alignment study to compare WorkKeys assessment content with the NAEP Mathematics and Reading items and frameworks (Dickinson et al., 2014). The study sought to expand upon the prior NAEP alignment studies by including additional WorkKeys assessments (Applied Technology and Graphic Literacy) that might align to NAEP Reading and Mathematics. Items and targets for instruction for Workplace Documents were paired with those for Graphic Literacy for analysis, while those for Applied Math were paired with those for Applied Technology. The results found that NAEP items did not represent the content domain, with 52% of mathematics targets and 72% of reading targets not matched to any NAEP item. Sixteen of the 24 content strands in the NAEP Mathematics framework and one of the three cognitive targets in the NAEP Reading framework could not be matched to any of the WorkKeys items.

The NAGB concluded that there was weak alignment between the WorkKeys and NAEP assessments, which was expected due to their differing purposes (i.e., NAEP is not a career or work assessment but rather a measure of academic skills). Based on the low level of alignment and the evidence ACT provided concerning the relationship between WorkKeys content and job preparedness, the study concluded that NAEP is not an appropriate measure of academic preparedness for job training (Dickinson et al., 2014).

In summary, the WorkKeys assessments show strong alignment with frameworks articulating skills pertinent to adult education and employability, and there is alignment across WorkKeys and other frameworks in terms of the various levels' complexity, breadth, and depth of competency. That is, not only are the same skills articulated, but as they increase in complexity in one framework, they similarly increase in complexity in WorkKeys. However, no strong alignment has been observed for NAEP, which is intuitive, given that it is a measure of academic preparedness rather than work readiness and the assessments have different use cases.

Evidence Based on Internal Structure

Differential Item Functioning

Some internal structure studies focus on whether particular items function differently across subgroups (e.g., racial/ethnic subgroups or gender subgroups). Differential item functioning (DIF) detects any statistical differences in item responses between subgroups assumed to have equal ability. When sufficient numbers of test takers are available, ACT uses DIF analyses to evaluate and flag operational items that could be unfair to any group of test takers. Items found to be fair in earlier qualitative reviews can still function differently for specific population

subgroups. ACT uses Mantel-Haenszel Delta (MHD) statistics (Dorans & Holland, 1993) to detect the existence of DIF in WorkKeys items. Items found to exceed critical values for DIF are further reviewed by content specialists and potentially also by external evaluators who have training and expertise in cultural anthropology or multicultural education. The results of this review may lead to the removal of one or more items from a form (ACT, 2025a).

ACT has conducted fairness reviews examining DIF in Applied Math, Graphic Literacy, and Workplace Documents. To examine item-level fairness, ACT conducted a preliminary DIF analysis on operational items (two forms for each construct). The findings indicated that the majority of the items in the investigated forms did not perform differently between female and male test takers, Black and White[4] test takers, or Hispanic/Latino and White test takers (ACT, 2023).

Demographic Group Analyses

Tables 4–6 present supplemental analyses of data comparing mean WorkKeys level scores by race/ethnicity, gender, and age. In all instances, mean scores were statistically different from one another, though the tests are likely overpowered given the large sample sizes. For all analyses, standardized effect sizes (*d*) are reported to provide insight regarding the magnitude of the effect (Cohen, 1977). Because the tables report the magnitude of the effect of the mean differences, the results can be viewed in terms of their practical significance in interpreting differences in level scores. The meaning of effect size varies by context, but the standard interpretation offered by Cohen (1988) is that .80 = large effect, .50 = moderate effect, and .20 = small effect.

Mean differences for ethnic groups on cognitive assessments are a common finding (Gottfredson, 1988; Roth et al., 2001; Ryan, 2001; Sackett & Wilk, 1994). Performance on the WorkKeys cognitive assessments is consistent with these findings.

Table 4. Results of *t*-Test and Descriptive Statistics for WorkKeys Mean-Level Scores by Race/Ethnicity

ACT WorkKeys assessment	White			Non-White			*t*	df	*d*
	n	*M*	*SD*	*n*	*M*	*SD*			
Applied Math	1,631,837	4.95	1.41	1,539,375	4.15	1.67	457.63*	3,015,975	.52
Graphic Literacy	1,427,486	4.06	.96	1,013,784	3.59	1.23	312.56*	2,370,323	.43
Workplace Documents	1,626,558	5.07	1.18	1,192,107	4.53	1.29	357.53*	3,098,849	.43

*$p < .001$.

Table 5. Results of *t*-Test and Descriptive Statistics for WorkKeys Mean-Level Scores by Gender

ACT WorkKeys assessment	Male			Female			*t*	df	*d*
	n	*M*	*SD*	*n*	*M*	*SD*			
Applied Math	1,651,372	4.70	1.61	1,483,814	4.43	1.58	153.25*	3,111,549	.17
Graphic Literacy	1,422,567	3.86	1.17	1,227,836	3.87	1.06	−5.43*	2,643,152	−.01
Workplace Documents	1,646,941	4.76	1.35	1,486,163	4.87	1.18	−74.29*	3,128,851	−.08

*$p < .001$.

Table 6. Results of *t*-Test and Descriptive Statistics for WorkKeys Mean-Level Scores by Age

ACT WorkKeys assessment	Under age 40			40 or older			*t*	df	*d*
	n	*M*	*SD*	*n*	*M*	*SD*			
Applied Math	2,899,385	4.56	1.60	507,016	4.57	1.54	−3.52*	713,234	.00
Graphic Literacy	2,408,959	3.88	1.11	511,006	3.73	1.17	86.27*	718,231	.14
Workplace Documents	2,895,981	4.78	1.27	506,027	5.07	1.25	−156.09*	700,574	−.24

*$p < .001$.

In addition to examining NCRC levels earned by race/ethnicity, gender, and age group, Allen (2025) recently examined NCRC levels earned by education group and employment status. With regard to those group differences, his primary takeaways were that a) performance increased consistently with education level and b) examinees who were college students, recent college graduates, or employed performed better than high school students, recent high school graduates, and those who were unemployed.

Evidence Based on Relations to Other Variables

Often, an important source of validity evidence is the relationship between test scores and external variables (AERA et al., 2014; EEOC, 1978; SIOP, 2018). This type of validity evidence includes evidence for convergent and discriminant validity. Convergent validity investigates the relationship between test scores and other measures designed to assess similar constructs. Discriminant validity, on the other hand, investigates the relationship between test scores and measures of different constructs. For example, a strong correlation[5] between the Applied Math assessment and another existing measure of applied mathematics would provide convergent evidence. Discriminant evidence would include studies that show Applied Math has a stronger relationship with an existing test of applied mathematics than it does with tests of unrelated constructs such as manual dexterity, leadership skills, or strength.

Another type of validity evidence based on relations to other variables is evidence related to test-criterion relationships. This speaks to how accurately test scores predict criterion performance, a measure of some outcome that is operationally distinct from the test. For example, if a self-reported measure of workplace performance predicted wages, it would be said to have test-criterion validity evidence. There are two main types of test-criterion validation research studies: concurrent and predictive. In a concurrent study, test scores and criterion information are collected at approximately the same time. In a predictive study, criterion information is collected some time after test scores are collected.

Convergent and Discriminant Validity Evidence

To support the validity of Workplace Documents test scores, ACT examined the relationship between Workplace Documents and the ACT English and reading tests, which measure the language skills identified as prerequisites to successful performance in entry-level college courses in English and reading (ACT, 2008a). ACT collected score data from two test administrations in a midwestern state, one from a sample of 121,304 students in spring 2002 and another from a sample of 122,820 students in spring 2003. Test takers who received higher scale scores on the ACT English and reading tests generally received higher level scores on Workplace Documents (ACT, 2008a). Using the same data, researchers also examined the relationship between WorkKeys Applied Math and the ACT mathematics test to provide construct validity evidence (ACT, 2008b). The ACT mathematics test measures the mathematical reasoning skills identified as prerequisite to successful performance in entry-level courses in college mathematics. The results indicate that Applied Math scores are highly correlated with ACT mathematics scores. In general, test takers who received higher scores on Applied Math also received higher scale scores on the ACT math test (ACT, 2008b).

In two distinct samples, Buchanan (2000; a study of incarcerated individuals) and Conway (2023a; a study of community college students) reported correlations between WorkKeys Applied Math, Graphic Literacy, and Workplace Documents scores and TABE (an assessment of adult basic and secondary education knowledge) Language, Reading, and Math scores. All the tests were expected to be positively correlated with one another given that all measure some aspects of intelligence, a unidimensional construct. This hypothesis was confirmed, thereby providing evidence of convergent validity. In terms of discriminant validity, results were mixed. In both studies, there was evidence of discriminant validity for Workplace Documents, as it correlated more highly with TABE Reading and Language than Math, but there was no strong evidence of discriminant validity for Applied Math. (There were no specific expected relationships between Graphic Literacy and the TABE scores.) Additional evidence of convergent validity was provided by Steedle and Hepburn (2020), who reported that Workplace Documents correlated .55 with Lexile reading scores in a study of the relationships between WorkKeys and outcomes for students in a youth manufacturing apprenticeship program. In this instance, the correlation .55 indicates a moderate-large association between Workplace Documents and Lexile reading scores.

Test-Criterion Validity Evidence

In addition to convergent and discriminant validity evidence, test-criterion relationship evidence is critical. Different outcomes or criteria may be relevant to different uses of the WorkKeys assessments and NCRC. For example, the WorkKeys assessments and NCRC are used for employment decisions where job performance is the primary outcome of interest. Job performance is typically measured via supervisor ratings of performance, but other criteria or employee outcomes such as attendance and safety incidents may be used in specific studies. When WorkKeys tests are used for employee development or the assessment of individual or group readiness, criteria other than job performance may be more relevant (e.g., individual earnings, employment, and training completion). Evidence that supports the use of WorkKeys or the NCRC to predict one outcome may not necessarily be generalized to other outcomes.

The correlations between the WorkKeys tests and job performance ratings provide criterion-related evidence for the validity of using WorkKeys in relation to a specific job. A number of studies have been conducted across a range of organizations to examine the relationship between WorkKeys cognitive test scores and employee job performance ratings. Sample sizes and correlations vary across studies of a wide spectrum of occupations and across the assessments. Early WorkKeys criterion validity studies relied on measures of job performance based on job- and company-specific task lists developed during the job profiling process. Studies conducted between 2006 and 2015 used the ACT Supervisor Survey or ACT WorkKeys Appraise, both of which rely on more generalized categories of job performance based on literature about common dimensions of job performance (ACT, 2015b).

Early Criterion Validation Studies

Between 1993 and 2015, numerous criterion validation studies were conducted on the WorkKeys assessments, resulting in unpublished manuscripts. These studies represent the first criterion validation studies conducted after WorkKeys was developed. Table 7 provides a

breakout of the number of unique studies by assessment and includes the range of sample sizes and correlations. The table also presents the relationship between various outcome measures and composite scores from Applied Math, Graphic Literacy, and Workplace Documents (LeFebvre, 2016). The results of the studies show a modest positive relationship with supervisor ratings of overall job performance; a positive relationship with education outcomes such as grade point average (GPA), course grades, and postsecondary persistence; and a negative relationship with safety incidents, customer complaints, absenteeism, and turnover (i.e., higher WorkKeys assessment scores are associated with lower rates of safety incidents etc.). For example, the composite Applied Math, Graphic Literacy, and Workplace Documents score has a median validity coefficient of .29 for supervisor-rated overall job performance. Interestingly, the predictive validity estimate of the Hogan Personality Inventory, which is among the most widely used and well-validated personality inventories for organizational applications, is .29 for job performance across job families (Hogan Assessments, 2017). Therefore, we can conclude that the WorkKeys assessments are as valid as other trusted and widely used assessments.

Table 7. Summary of ACT WorkKeys Criterion Validation Studies by Assessment and Outcome: 1993–2015

WorkKeys assessment	Outcomes	No. of studies	*N*		Range of validity coefficients		
			Min.	Max.	Min.	Med.	Max.
Applied Math (AM)	Career tech course grades	1	2,162	2,162	.21	.21	.21
	Postsecondary GPA	1	1,246	1,246	.28	.28	.28
	Overall job performance—supervisor ratings	13	13	165	−.23	.12	.41
Graphic Literacy (GL)	Career tech course grades	1	1,216	1,216	.21	.21	.21
	HRIS data—turnover	1	96	96	−.33	−.33	−.33
	HRIS data—absenteeism	1	96	96	−.22	−.22	−.22
	HRIS data—safety incidents	1	96	96	−.11	−.11	−.11
	HRIS data—customer complaints	1	96	96	−.11	−.11	−.11
	Overall job performance—supervisor ratings	14	13	314	−.51	.16	.32
Workplace Documents (WD)	Career tech course grades	1	2,223	2,223	.22	.22	.22
	Postsecondary GPA	1	1,251	1,251	.25	.25	.25
	HRIS data—turnover	1	96	96	.12	.12	.12
	HRIS data—absenteeism	1	96	96	−.13	−.13	−.13
	HRIS data—safety incidents	1	96	96	−.15	−.15	−.15
	HRIS data—customer complaints	1	96	96	−.24	−.24	−.24
	Overall job performance—supervisor ratings	16	10	314	−.32	.20	.86
Composite AM and WD	Postsecondary GPA	1	10,744	10,744	.30	.30	.30
	College persistence	1	277,631	277,631	.23	.23	.23
Composite AM, GLA, and WD	Overall job performance—supervisor ratings	3	68	951	.29	.29	.29
	Career tech course grades	1	951	951	.25	.25	.25

Note. Of the many dimensions of job performance studied, only the overall job performance correlations are reported in this table for summary purposes. HRIS = Human Resource Information System data collected by employer

Recent Criterion Validation Studies

Since 2015, ACT and others have published numerous reports documenting the criterion validity evidence of the NCRC. First, the NCRC has been linked to education-related outcomes such as ACT scores, GPA, program curriculum, and program completion. In addition, the NCRC and its assessments have been associated with workforce-related outcomes such as credentialing, employment, retention, wages, and absences.

Education Outcomes

One study of nearly 79,000 students who took both the ACT and all three WorkKeys NCRC assessments (Radunzel & Fang, 2018) determined that as ACT Composite scores increase, a larger percentage of students are predicted to reach the Gold or Platinum level on the NCRC. Based on binary logistic regression prediction models, Radunzel and Fang reported that students who earn an ACT Composite score of less than 13 are unlikely to obtain an NCRC, students with scores between 13 and 16 are most likely to obtain a Bronze NCRC, students with scores between 17 and 21 are most likely to obtain a Silver NCRC, students with scores between 22 and 26 are most likely to obtain a Gold NCRC, and students with scores between 27 and 36 are most likely to obtain a Platinum NCRC. Allen et al. (2016) reported a similar pattern of findings. More recently, Allen (in press) examined the relationship between ACT and NCRC scores but in reverse: he used the NCRC assessments as the predictor variables and ACT Composite scores as the outcome variable in a study of nearly 570,000 students who took the ACT and WorkKeys during the same academic year and within close time proximity. He reported that 65% of the variance in ACT Composite scores could be accounted for by Applied Math, Graphic Literacy, and Workplace Documents scores. Furthermore, he reported the following predicted ACT Composite score range for each NCRC level: non-qualifier = 12–14, Bronze = 13–16, Silver = 15–20, Gold = 19–24, and Platinum = 23–29.

In addition to being linked with ACT scores, the NCRC has been linked with postsecondary GPA. In one study stemming from a research partnership between ACT and the Mississippi Community College Board, students with higher scores on the NCRC and its component assessments tended to have higher GPAs (Conway, 2023a). Correlations between the assessments and GPA were as follows: Applied Math = .16, Graphic Literacy = .14, Workplace Documents = .12, and NCRC = .17. These correlations were higher, reaching .40 for each component assessment and .41 for the NCRC, in a study of first-year college GPA of students within the Arkansas higher education system (Conway, 2023b). A link between the NCRC and GPA was also reported in a study of students at the Cincinnati State Technical and Community College, in which Steedle et al. (2017) found that students without an NCRC had an average GPA of 2.2, while those who obtained an NCRC had an average GPA of 3.1. Steedle and LeFebvre (2018) provided a more fine-grained account of cumulative GPA across NCRC levels in a study of postsecondary students, noting that each increase in NCRC level was associated with an approximate .22 increase in GPA on a 4.0 scale. For non-qualifiers, the average GPA was 2.56, and average GPAs were 2.74, 2.99, and 3.21 respectively for Bronze, Silver, and Gold/Platinum levels (combined due to a small number of Platinum earners). Conway (2023b) documented the same stair-step pattern of first-year GPAs in the study of students in Arkansas,

with students of different NCRC levels achieving the following average GPAs: non-qualifiers = 2.81, Bronze = 2.98, Silver = 3.29, Gold = 3.52, and Platinum = 3.63.

Course-taking patterns differ across individuals at varying NCRC levels as well. Students in the youth manufacturing apprenticeship program referenced above (Steedle & Hepburn, 2020) with higher Applied Math scores were more likely to have taken higher-level math courses during their senior year. Students whose highest course was Technical Math scored an average of 3.6 points higher on Applied Math than students whose highest course was Math for the Trades. This corresponds to a 1.0 standard deviation difference on the Applied Math score scale, which is considered a large effect. Another study (Williams, 2015) showed that community college students with lower NCRC scores tended to be placed in remedial versus regular math and English courses.

Several studies have examined the association between the NCRC and degree attainment. Data show that the NCRC is related to earning one's high school equivalency (Conway, 2023a). Students who earned their NCRC were three times as likely to earn their high school equivalency (75% vs. 25%). While just 25% of non-qualifiers earned their high school equivalency, 96% of individuals scoring at the Platinum level did. What is more, the NCRC showed incremental validity in predicting high school equivalency attainment over and above the three TABE assessments mentioned above (an additional 5%). In a second study published by ACT (Steedle & LeFebvre, 2018a), NCRC levels were associated with degree attainment. For example, during the period of this particular study, between the Bronze, Silver, and Gold/Platinum levels, the percentage of individuals earning an associate's degree increased from 66% to 71% to 76%. Two third-party reports of the criterion validity related to program completion should be noted as well. Among students within a postsecondary career and technical education program, Applied Math scores were significantly higher for students who completed their program than for students who did not (d = .16; Bowers, 2024). The same was true for Graphic Literacy (d = .12) but not for Workplace Documents (d = .03). In contrast, however, Williams (2015) did not find a relationship between NCRC levels and program completion among career technical students enrolled at a rural Mississippi community college.

Workforce Outcomes

The body of literature in the realm of education outcomes provides strong support for the WorkKeys assessments and the NCRC. We turn now to workforce outcomes, including credentialing, employment, retention, wages, and absences.

In preparation for or while participating in apprenticeship programs or technical programs, students must often sit for industry credential exams, such as Certified Production Technician (CPT) certificates, and NCRC levels have been linked with success on these exams. For example, apprentices demonstrating greater foundational workplace skills on WorkKeys exhibited higher average achievement on CPT exams measuring manufacturing knowledge and skills (Steedle & Marshall, 2019). Similarly, Steedle and Hepburn (2020) reported slight positive correlations between the three NCRC assessments and the CPT exams. For Applied Math, correlations reached .24 with CPT Safety; for Graphic Literacy, correlations reached .19 with

CPT Safety; and for Workplace Documents, correlations reached .25 with CPT Manufacturing Processes and Production. Bowers (2024) compared the NCRC scores of students who earned at least one industry-recognized credential with the scores of those who earned none. On average, Applied Math (d = .34), Graphic Literacy (d = .45), and Workplace Documents (d = .17) scores were higher for the former group.

With regards to the NCRC's association with employment, in one study of community college students (Steedle et al., 2017), more students who earned an NCRC (70%) became employed compared with students who did not earn an NCRC (57%). Steedle and LeFebvre (2018a) also examined data indicating that higher NCRC levels were associated with a greater likelihood of employment. These findings counter those of Williams (2015), who found no association between NCRC levels and employment status. She noted, however, that there was a large number of students for whom employment status was unknown, and for the remainder, the number of employed students was very high. Related, one study suggests that duration of unemployment is unrelated to NCRC levels (Steedle & LeFebvre, 2018). The authors noted, though, that a number of factors could mask the effect of earning higher NCRC levels. For instance, the jobs that would attract those with higher NCRC levels may be scarcer; may have lengthier application, interview, and hiring processes; or simply may not recognize the NCRC.

The WorkKeys assessments' link with retention has also been queried. Hendrick (2006) compared the retention of employees hired with and without WorkKeys scores and determined that the use of WorkKeys in the hiring process increased employee retention. Eighty-seven percent of employees hired using WorkKeys were employed for longer than 12 months, while less than 80% of those hired without WorkKeys scores were employed for longer than 12 months. Moreover, Hendrick showed that the percentage of employees who persisted in their jobs increased as scores increased. Finally, she examined whether it was beneficial to use additional WorkKeys assessments such as Applied Tech and Business Writing rather than just the three that make up the NCRC. She determined that there was a difference in job retention between employees hired using tests beyond the NCRC trio and employees hired using only the three NCRC assessments, with a greater number of assessments leading to greater retention rates.

In addition to research about industry-related credentials, employment, and retention, there is research suggesting that the NCRC is associated with income. Steedle and LeFebvre (2018a) obtained quarterly wage records from a database 9 months before and 15 months after examinees took WorkKeys. For the more than 18,000 examinees with post-test records, average quarterly wages increased steadily from $3,600 for non-qualifiers to $6,549 for Gold/Platinum NCRC earners. In a separate study (Steedle et al., 2017), when comparing those who earned any NCRC with non-qualifiers, the researchers found that average weekly wages increased $96 for earners versus $54 for non-qualifiers over time. Finally, in their second study, Steedle and LeFebvre (2018b) reported findings to support the claim that earning higher NCRCs can help individuals secure higher incomes in both the short and long term. ACT delivered a list of a random sample of 50,000 WorkKeys examinees to Equifax, who matched it to income records from 2011 to 2016. The study found that, for high school testers, median incomes were similar for all NCRC levels between 2011 and 2013, but they diverged from 2014

to 2016, with incomes increasing at faster rates for those with higher NCRC levels. In addition, for adult testers with high school diplomas, there was a substantial jump over the study period in median income for those with a Gold/Platinum NCRC, much different from the national average, which was flat during the study period.

Finally, there is some indication that NCRC levels are linked to absences. In a study of individuals in a youth manufacturing apprenticeship program (Steedle & Hepburn, 2020), the median number of absences during the senior year for non-qualifiers was nearly 16, while the median number for those who earned an NCRC was just 5. Furthermore, albeit a small effect, students with higher NCRC levels had fewer average absences than those with lower levels (5.1 for Bronze, 4.7 for Silver, 4.5 for Gold, and 4.5 for Platinum).

Conclusion

The ACT WorkKeys assessments originally relied solely on evidence based on test content. More recently, evidence based on internal structure and relations to other variables has accumulated. Research has demonstrated associations with important outcomes in educational, employment training, and work environments. These additional pieces of evidence have bolstered the validity argument for the NCRC and its associated assessments. Validation is an ongoing process and is often the joint responsibility of the test developer and organizations using the assessments. ACT remains committed to providing multiple sources of evidence to support the interpretative arguments for and intended uses of these WorkKeys assessments and the NCRC.

References

ACT. (1987). *Study power* [Series].

ACT. (1992). *A strategic plan for the development and implementation of the WorkKeys system* [Unpublished report].

ACT. (2008a). *WorkKeys Reading for Information technical bulletin*.

ACT. (2008b). *WorkKeys Applied Mathematics technical bulletin*.

ACT. (2009). *WorkKeys Talent assessment user and technical guide.* https://www.act.org/content/dam/act/unsecured/documents/WK-Talent-Assessment-User-and-Technical-Guide.pdf

ACT. (2014a). *ACT WorkKeys assessments and the National Network of Business and Industry Associations' common employability skills framework* [Alignment study]. https://www.act.org/content/dam/act/unsecured/documents/WK-Study-NNBIA-Skills-Framework.pdf

ACT. (2014b). *JobPro gap analysis* [Unpublished report].

ACT. (2021). *ACT WorkKeys NCRC crosswalk to College and Career Readiness Standards for Adult Education*. https://www.workreadycommunities.org/files/pdf/WorkKeys%20NCRC%20Crosswalk%20to%20CCRSAE%20Report%20&%20Appendices%201.6.22.pdf

ACT. (2024). *ACT WorkKeys Essential Skills technical manual.* https://www.act.org/content/dam/act/unsecured/documents/pdfs/ACT-WorkKeys-Essential-Skills-Technical-Manual.pdf

ACT. (2025a). *WorkKeys NCRC assessments technical manual*. https://www.act.org/content/dam/act/unsecured/documents/ACT-workkeys-NCRC-technical-manual.pdf

ACT. (2025b). *ACT WorkKeys accessibility supports guide.* https://www.act.org/content/dam/act/unsecured/documents/WorkKeysAccessibilitySupportsGuide.pdf

ACT. (2025c). *ACT technical manual*. https://www.act.org/content/dam/act/unsecured/documents/ACT_Technical_Manual.pdf

Agency for Instructional Technology. (1989). *Workplace readiness: Education for employment, personal behavior, group effectiveness and problem solving for a changing workplace* [Unpublished report].

Allen, J. (2025). *A summary of WorkKeys participation and performance in the United States: 2023–2024* [Unpublished report]. ACT.

Allen, J. (in press). *Predicting ACT scores from NCRC performance*. ACT.

Allen, J., LeFebvre, M., & Mattern, K. (2016). *Predicted attainment of ACT National Career Readiness Certificates for the 2016 ACT-tested high school cohort* (Technical Brief R1602). ACT. https://www.act.org/content/dam/act/unsecured/documents/Predicted-Attainment-of-ACT-National-Career-Readiness-Certificates-for-the-2016-ACT-tested-High-School-Cohort.pdf

American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. American Educational Research Association. https://www.testingstandards.net/open-access-files.html

Bailey, L. J. (1990). *Working: Skills for a new age*. Delmar.

Bowers, B. (2024). *Efficacy of the WorkKeys ACT assessments in predicting student success and postsecondary career technical education and industry recognized credential attainment* [Master's thesis, Shawnee State University]. Digital Commons. https://digitalcommons.shawnee.edu/cgi/viewcontent.cgi?article=1079&context=math_etd

Buchanan, B. L. (2000). *A comparative study of the tests of adult basic education and Work Keys with an incarcerated population* [Doctoral dissertation, Texas A&M University]. ProQuest Dissertations & Theses. https://www.proquest.com/openview/181d11f70039a1e7c2378d5e2e305efd/1?pq-origsite=gscholar&cbl=18750&diss=y

Carnevale, A. P., Gainer, L. J., & Meltzer, A. S. (1990). *Workplace basics: The essential skills employers want*. Jossey-Bass. https://files.eric.ed.gov/fulltext/ED319979.pdf

Cascio, W. F. (1982). *Applied psychology in personnel management* (2nd ed.). Reston Publishing.

Center for Occupational Research and Development. (1990). *CORD applied mathematics* [Curriculum].

Clark, H. (2015). *Building a common language for career readiness and success: A foundational competency framework for employers and educators* (Working Paper Series WP-2015-02). ACT. https://www.act.org/content/dam/act/unsecured/documents/WP-2015-02-Building-a-Common-Language-for-Career-Readiness-and-Success.pdf

Cohen, J. (1977). *Statistical power analysis for the behavioral sciences* (Rev. ed.). Routledge.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.

Conover Company. (1991). *Workplace literacy system* [Brochure].

Conway, J. (2023a). *Evaluating the validity of WorkKeys in the Mississippi community college system* (Technical Brief R2259). ACT. https://www.act.org/content/dam/act/unsecured/documents/R2259-Evaluating-WorkKeys-Validity-in-the-MS-Community-College-System-03-2023.pdf

Conway, J. (2023b). *Criterion-related and differential validity of ACT WorkKeys in Arkansas* (Technical Brief R2131). ACT. https://www.act.org/content/dam/act/unsecured/documents/R2131-Criterion-Related-and-Differential-Validity-of-ACT-WorkKeys-in-Arkansas-07-2023.pdf

Dickinson, E. R., Smith, E., Deatz, R., Thacker, A. A., Sinclair, A. L., & Johnston-Fisher, J. (2014). *The content alignment between the NAEP and WorkKeys assessments*. NAGB. https://www.nagb.gov/content/dam/nagb/en/documents/what-we-do/preparedness-research/content-alignment/naep_workkeys_final.pdf

Dorans, N. J., & Holland, P. W. (1993). DIF detection and description: Mantel Haenszel and standardization. In P. W. Holland & H. Wainer (Eds.), *Differential item functioning* (pp. 35–66). Lawrence Erlbaum Associates.

Dunnette, M. D., & Hough, L. M. (Eds.). (1990). *Handbook of industrial and organizational psychology* (2nd ed., Vol. 1). Consulting Psychologists Press.

Electronic Selection Systems Corporation. (1992). *AccuVision systems for personnel selection and development* [Assessment].

Equal Employment Opportunity Commission. (1978). *Section 60-3, Uniform Guidelines on Employee Selection Procedures (1978); 43 FR 38295 (August 25, 1978)*. EEOC 29 CFR Part 1607. https://www.uniformguidelines.com/uniform-guidelines.html

Gottfredson, L. S. (1988). Reconsidering fairness: A matter of social and ethical priorities. *Journal of Vocational Behavior*, *33*(3), 293–319. https://doi.org/10.1016/0001-8791(88)90041-3

Greenan, J. P. (1983). *Identification of generalizable skills in secondary vocational programs* [Executive summary]. Illinois State Board of Education. https://files.eric.ed.gov/fulltext/ED233223.pdf

Guttman, L. (1950). The basis for scalogram analysis. In S. A. Stouffer, L. Guttman, E. A. Suchman, P. A. Lazarsfeld, S. A. Star, & J. A. Clausen (Eds.), *Measurement and prediction* (pp. 60–90). Princeton University Press.

Hendrick, R. Z. (2006). *Evaluating Work Keys profiling as a pre-employment assessment tool to increase employee retention.* [Doctoral dissertation, Old Dominion University]. Digital Commons. https://digitalcommons.odu.edu/cgi/viewcontent.cgi?article=1186&context=efl_etds

Hill, C., & Lyons, D. (2017). *ACT WorkKeys NCRC crosswalk to beta Connecting Credential Framework*. ACT. https://www.act.org/content/dam/act/unsecured/documents/NCRC-Crosswalk-Research-2017-09.pdf

Hogan Assessments. (2017, May 30). *A quick and dirty guide to validity and reliability.* https://www.hoganassessments.com/blog/quick-dirty-guide-validity-reliability/

Kane, M., Berryman, S., Goslin, D., & Meltzer, A. (1990). *Identifying and describing the skills required by work*. Secretary's Commission on Achieving Necessary Skills; Pelavin Associates. https://babel.hathitrust.org/cgi/pt?id=coo.31924062239557&seq=3

Kolen, M. J., & Brennan, R. L. (2004). *Test equating, scaling, and linking: Methods and practices* (2nd ed.). Springer.

LeFebvre, M. (2016). *A summary of ACT WorkKeys validation research* (Research Report 2016-4). ACT. https://www.act.org/content/dam/act/unsecured/documents/5350-Research-Report-2016-4-A-Summary-of-ACT-WorkKeys-Validation-Research.pdf

Lord, F. M. (1980). *Applications of item response theory to practical test problems*. Lawrence Erlbaum Associates.

McLarty, J. R. (1992, August). *WorkKeys: Developing the assessments*. In J. D. West (Chair), *WorkKeys supporting the transition from school to work* [Symposium]. Meeting of the American Psychological Association, Washington, DC, United States.

McLarty, J. R., & Vansickle, T. R. (1997). Assessing employability skills: The Work Keys system. In H. F. O'Neil, Jr. (Ed.), *Workforce readiness: Competencies and assessment* (pp. 293–325). Lawrence Erlbaum Associates.

National Assessment Governing Board. (2010a). *The alignment of the NAEP grade 12 Mathematics assessment and the ACT WorkKeys Applied Mathematics assessment*. NAGB; ACT. https://www.nagb.gov/content/dam/nagb/en/documents/what-we-do/preparedness-research/content-alignment/WorkKeys-NAEP_Math_Content_Comparison.pdf

National Assessment Governing Board. (2010b). *The alignment of the NAEP grade 12 Reading assessment and the WorkKeys Reading for Information assessment*. NAGB; ACT. https://www.nagb.gov/content/dam/nagb/en/documents/what-we-do/preparedness-research/content-alignment/WorkKeys-NAEP_Reading_Content_Comparison.pdf

Radunzel, J., & Fang Y. (2018). *Updating the progress toward the ACT National Career Readiness Certificate indicator* (Technical Brief R1712). ACT. https://www.act.org/content/dam/act/unsecured/documents/R1712-ACT-progress-toward-NCRC.pdf

Roth, P. L., Bevier, C. A., Bobko, P., Switzer, F. S., III, & Tyler, P. (2001). Ethnic group differences in cognitive ability in employment and educational settings: A meta-analysis. *Personnel Psychology*, *54*(2), 297–330. https://doi.org/10.1111/j.1744-6570.2001.tb00094.x

Ryan, A. M. (2001). Explaining the black-white test score gap: The role of test perceptions. *Human Performance*, *14*(1), 45–75. https://doi.org/10.1207/S15327043HUP1401_04

Sackett, P. R., & Wilk, S. L., (1994). Within-group norming and other forms of score adjustment in preemployment testing. *American Psychologist*, *49*(11), 929–954. https://doi.org/10.1037/0003-066X.49.11.929

Schulz, E. M., Kolen, M. J., & Nicewander, W. A. (1997). *A study of modified-Guttman and IRT-based level scoring procedures for Work Keys assessments* (Research Report 97-7). ACT. https://www.act.org/content/dam/act/unsecured/documents/ACT_RR97-07.pdf

Society for Industrial and Organizational Psychology. (2018). *Principles for the validation and use of personnel selection procedures* (5th ed.). American Psychological Association. https://www.apa.org/ed/accreditation/personnel-selection-procedures.pdf

Steedle, J., & Hepburn, A. (2020). *The ACT WorkKeys NCRC as an indicator of skills needed for success in work-based learning* (Technical Brief R1796). ACT. https://www.act.org/content/dam/act/unsecured/documents/2022/R1796-gps-work-based-learning-2020-01.pdf

Steedle, J., & LeFebvre, M. (2018a). *Evaluating the ACT NCRC as a signal of the skills needed for labor market and educational success in Missouri* (Technical Brief R1728). ACT. https://www.workreadycommunities.org/files/pdf/R1728-ncrc-mo-success-2018-11.pdf

Steedle, J., & Marshall, S. K. (2019). *The ACT WorkKeys National Career Readiness Certificate: A foundation for stackable credentials in manufacturing* (Data Byte R1738). ACT. https://www.act.org/content/dam/act/unsecured/documents/R1738-ncrc-sworwib-2019-01.pdf

Steedle, J., Ndum, E., & Mattern, K. (2017). *The ACT National Career Readiness Certificate as a predictor of academic and workforce outcomes* (Technical Brief 1663). ACT. https://www.act.org/content/dam/act/unsecured/documents/pdfs/R1663-act-ncrc-outcomes-predictor-2017-12.pdf

Tatzel, M., & Lamdin, L. (1975). *Interpersonal learning in an academic setting: Theory and practice* (CAEL Institutional Report No. 2). ETS.

Vosper, L., Bolin, B., Byrd, K., & Wineinger, B. (2005). *The Career Readiness Certificate: Strategies toward a common goal*. Career Readiness Certificate Consortium. Retrieved March 1, 2012, from ncrcadvocates.org (site discontinued)

Williams, L. K. (2015). *ACT WorkKeys as an indicator of academic success* [Doctoral dissertation, Mississippi State University]. Scholars Junction. https://scholarsjunction.msstate.edu/cgi/viewcontent.cgi?article=1368&context=td

Notes

1. Reading, math, and locating information in graphics are included in the NCRC based on evidence that the majority of profiled careers and career clusters require these skills. However, organizations employing the NCRC for pre-employment screening or selection decisions should verify that these skills are important for their positions or careers. This can be done using a variety of methods and does not mandate a formal job analysis or job profiling study.

2. Other charter members included the American Association of Community and Junior Colleges, the National Association of State Directors of Vocational Technical Education Consortium, and the National Association of Secondary School Principals.

3. Number Properties and Operations; Measurement; Geometry; Data Analysis, Statistics, and Probability; and Algebra

4. The category White includes only test takers who do not also identify as Hispanic/Latino.

5. Often, we evaluate such validity evidence with correlation coefficients, which can range from −1.0 to +1.0, both of which indicate a perfect relationship between two variables. Correlations at or near 0 indicate no association between two variables.

Appendix

The following studies were included in the Table 7 summary.

ACT. (2010). *WorkKeys value across the pipeline* [Unpublished report].

ACT. (2011). *E. & J. Gallo: WorkKeys—uncorking an employee's potential* [Unpublished report].

ACT. (2012). *WorkKeys American Health Care Association study* [Unpublished report].

ACT. (2013a). *WorkKeys concurrent validation and talent benchmark study for a health care organization* [Unpublished report].

ACT. (2013b). *WorkKeys as a predictor of critical thinking* [Unpublished report].

ACT. (2013c). *Ohio MSSC CPT and NCRC: A partnership for industry-recognized credentialing*. https://www.act.org/content/dam/act/unsecured/documents/Case-Study-MSSC_Research_Brief.pdf

ACT. (2014a). *WorkKeys concurrent validation and talent benchmark study for a health care organization* [Unpublished report]

ACT. (2014b). *City of Albuquerque ACT WorkKeys validation study and comparison study* [Unpublished report].

Iowa Workforce Development. (2012). *Skilled Iowa: Iowa NCRC 2012*. https://web.archive.org/web/20160815195037/http://portal.iowaworkforce.org/SkilledIowaLinks/NCRC2012Report.pdf

Mayo, M. J. (2012). *Evaluation metrics, New Options New Mexico, 2011–2012* [Unpublished manuscript].

Partners for a Competitive Workforce. (2012). Work readiness collaborative. https://web.archive.org/web/20120809233325/www.competitiveworkforce.com/Work-Readiness.html

Schmidt, F., & Sharf, J. C. (2010). *Review of ACT's WorkKeys program relative to the Uniform Guidelines and more current professional standards* [Unpublished report].

Workforce Investment Board of Southwest Missouri. (2013). *Average earnings, employment, & retention by national career readiness certificate & education levels*.

Zimmer, T. E. (2012). Assessment of the ACT WorkKeys job skills test. *InContext*, *13*(6). https://www.incontext.indiana.edu/2012/nov-dec/article3.asp

ABOUT ACT

ACT is transforming college and career readiness pathways so that everyone can discover and fulfill their potential. Grounded in more than 65 years of research, ACT's learning resources, assessments, research, and work-ready credentials are trusted by students, job seekers, educators, schools, government agencies, and employers in the U.S. and around the world to help people achieve their education and career goals at every stage of life. Visit us at www.act.org.