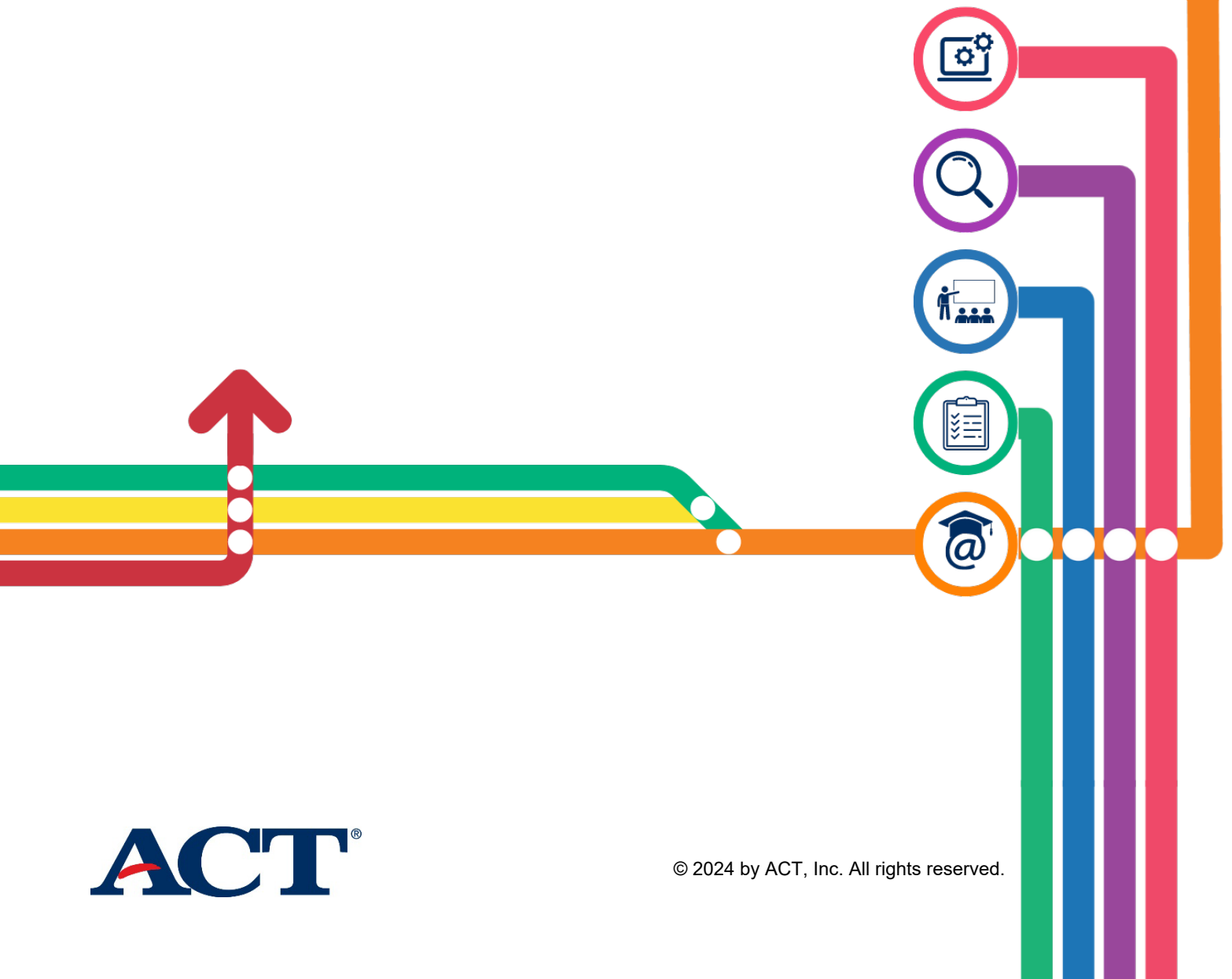


ACT[®] WorkKeys[®] Essential Skills Technical Manual

February 2024



Overview of the Technical Manual

This technical manual provides a brief overview of ACT® WorkKeys® Essential Skills, an assessment designed to measure six essential skills in employees within the workplace and potential employees looking to enter the workforce. The manual contains an overview of the skills assessed by WorkKeys Essential Skills, the two item types used to assess these skills, and the procedure for incorporating the results from the different item types into a single aggregate score. In addition, it provides a short description of the initial reliability and validity evidence for the assessment. Immediately below is a summary of the main points and findings. Subsequent sections of this manual provide additional detail on each of these points.

- WorkKeys Essential Skills assesses five essential skills aligned with the Big Five model of personality: work ethic, collaboration, resilience, creativity, and leadership. Additionally, the assessment measures integrity, which is aligned with the HEXACO model of personality.
- The six essential skills are assessed with two item types—self-report Likert items and situational judgment test (SJT) items—which are combined to yield one aggregate score per skill.
- A pilot study was conducted with nearly 1,800 employed adults to provide construct-related validity evidence. The study found that each skill maps onto its respective Big Five personality domain, while the integrity facet aligns with the honesty-humility construct from the HEXACO framework.
- An additional pilot study of 173 employed adults collected supervisor ratings of job performance. This criterion-related validity evidence supports the ability of WorkKeys Essential Skills scores to predict job performance.

Recommended Uses

- Employee and prospective employee development (alongside the ACT WorkKeys Essential Skills Coaching Tool)
- Applicant selection

Chapter 1: Introduction to WorkKeys Essential Skills

ACT® WorkKeys® Essential Skills is a comprehensive assessment system designed to measure six essential skills. The online assessment provides employees and organizations with a holistic picture of employees' strengths and areas in need of improvement. The assessment system includes individual-level reports that provide feedback on each skill, as well as roster reports (csv files containing all examinee scores, which can be sorted as needed or merged with an organization's data). The system also contains coaching activities to help employees improve their skills in each of the six skill areas. A complete curriculum that helps employees work on developing their skills can be purchased separately.

This technical manual details the empirical basis for the WorkKeys Essential Skills framework, the skills assessed by the assessment, the two item types used, and the procedure used to incorporate the item types into an aggregate score for each of the six skill areas. In addition, evidence of reliability and validity is provided, as are normative data.

Chapter 2: WorkKeys Essential Skills Background and Skills Assessed

What Are Essential Skills?

Essential skills, distinct from cognitive factors or intelligence, “include interpersonal, self-regulatory, and task-related behaviors important for adaptation to and successful performance in education and workplace settings” (Camara et al., 2015, p. vi). These skills have been shown to predict important life outcomes such as academic success, occupational attainment, and longevity (ACT, 2021; Roberts, Kuncel, et al., 2007). Furthermore, seven of the ten most requested skills in job postings are essential skills, employers seek essential skills 3.8 times more frequently than they do the top five technical skills (Cole et al., 2021), and essential skills are expected to grow in importance by 24% by 2030 (Bughin et al., 2018). Several alternative terms exist for these skills, including *durable skills*, *soft skills*, *psychosocial skills*, *personal skills/qualities*, and *character strengths* (Kyllonen et al., 2014). In the interest of clarity, we will use the term *essential skills*, as this term is commonly used within organizations when referring to noncognitive skills.

WorkKeys Essential Skills Organizing Framework: The Five-Factor Model of Personality

Five-Factor Model Overview

WorkKeys Essential Skills uses the five-factor model of personality (FFM; Digman, 1990) as its organizing framework. While many frameworks exist to explain and measure essential skills, the FFM of personality has emerged as the front-runner. The FFM framework was chosen because it effectively organizes the many essential skills that are critical for workplace success. It has been replicated in many countries and languages (McCrae et al., 2005), and a large body of psychological research has established the five factors’ importance for attaining many valued outcomes in education, the workplace, and life (Mammadov, 2021; Poropat, 2009; Roberts, Kuncel, et al., 2007; Zell & Lesick, 2021). Nearly all essential skills can be aligned with the FFM personality factors (Goldberg, 1993; Walton et al., 2021). Having amassed decades’ worth of information on this framework, researchers and practitioners both support the use of the FFM within the workplace. For these reasons, the development team utilized the FFM when developing WorkKeys Essential Skills.

The FFM is composed of five broad factors that capture the majority of personality trait variance. Guided by the lexical hypothesis, which assumes that important individual differences will become encoded into language as single terms (Goldberg, 1993), Allport and Odbert (1936) searched Webster’s 1925 *New International Dictionary* for English words describing human characteristics. In total, they selected roughly 18,000 English words and classified 4,500 as descriptions of stable personal traits. Cattell (1943) applied factor analytic procedures to reduce the massive list of traits by analyzing the underlying patterns among them. Subsequent efforts to reduce the list resulted in five replicable factors (Fiske, 1949; Norman, 1963; Tupes &

Christal, 1961). These constructs are commonly referred to as the Big Five and comprise conscientiousness, agreeableness, emotional stability (often referred to by its opposite pole, neuroticism), openness to experience, and extraversion (see de Raad & Mlačić, 2015, for a complete history).

Conscientiousness primarily describes the tendency to control impulses in an effort to achieve one's goals. People high on the conscientiousness factor can be described as dependable, organized, reliable, rule following, and hardworking. Agreeableness is associated with a prosocial and communal orientation toward others. People who are very agreeable tend to be tolerant, cooperative, friendly, and helpful, as well as trustworthy and trusting. Emotional stability concerns the ability to successfully cope with stressful situations and minimize the experience of negative emotions such as stress and anxiety. People high in emotional stability are poised and composed and can regulate their emotions under stressful or uncertain conditions. Openness to experience can be described as a person's interest in and acceptance of the unfamiliar. People high in openness to experience tend to be curious, inquisitive, thoughtful, and open-minded, and they often look for opportunities to express themselves creatively. Extraversion describes a person's desire for active social interaction. People who reflect high levels of extraversion are usually sociable, assertive, enthusiastic, and influential. They often enjoy leadership positions, feel comfortable expressing unpopular opinions, and experience positive emotionality. Together, these factors capture a great deal of information about an individual's overall personality (for a review of each factor, see John et al., 2008).

Although the Big Five were first discovered in the English language, studies in other languages and cultures, either involving the full psycholexical approach or applying translations of established FFM inventories, have resulted in the same five factors (see, e.g., McCrae et al., 2005; Schmitt et al., 2007). As one example, McCrae and colleagues (2005) documented the replication of the FFM in nearly 50 cultures on six continents. This evidence, as well as other research evidence from a wide variety of countries, cultures, and languages around the world, supports the FFM's universality.

In addition to having empirical support, the FFM taxonomy has also been widely adopted because it optimizes bandwidth and fidelity (Cronbach & Gleser, 1965), explaining a large amount of information while simultaneously retaining nuanced depictions of individual differences (Soto & John, 2014). Given that the Big Five are broad and multifaceted, this model can be used to integrate the plethora of essential skill terms discussed in the literature or assessed in practice. That is, it is likely that most or all essential skills can be mapped to the Big Five, which can be considered something of a "Rosetta stone" for understanding these skills (Roberts et al., 2015). Using the Big Five, for example, we can take constructs expressed as time management in one framework and responsibility in another and understand their connectedness as manifestations of conscientiousness.

HEXACO Model Overview

Although the Big Five framework is the most commonly accepted personality taxonomy, there is some research support for the HEXACO model. The HEXACO is a six-factor model of personality developed using factor analyses of personality descriptors from seven languages and cultures around the world. The HEXACO factors are honesty-humility (H), emotionality (E), extraversion (eX), agreeableness (A), conscientiousness (C), and openness to experience (O; Ashton et al., 2004). Thus, the HEXACO is similar to the Big Five, with the largest exception being the honesty-humility factor (although other minor differences do exist); honesty-humility typically falls under the agreeableness factor in the Big Five. Given that this factor shows value in predicting workforce-related outcomes of interest, such as job performance (Johnson et al., 2011) and self-reported workplace integrity and delinquency (Lee et al., 2005; Lee et al., 2009), an integrity skill was included in WorkKeys Essential Skills. This skill was designed to align with the HEXACO honesty-humility construct.

Essential Skills' Associations With Work Outcomes

As workplaces begin to focus on essential skills to meet the demands of increasingly adaptive and collaborative work environments, it is important to consider what types of essential skills can best predict important work outcomes. Industrial/organizational psychology provides substantial evidence concerning the role of essential skills (typically in the form of Big Five/HEXACO traits) in predicting important workplace outcomes, including task performance, engaging in appropriate and ethical workplace conduct, use of interpersonal skills (e.g., leadership, teamwork), work satisfaction, and perceived work stress (e.g., Judge et al., 1999; Lindqvist & Vestman, 2011; Ones et al., 1993; Roberts, Harms, et al., 2007; Schmidt & Hunter, 1998; van Iddekinge et al., 2012). Meta-analyses, which quantitatively aggregate and summarize findings from many individual research studies, have indicated that essential skills predict job performance (Barrick & Mount, 1991; Barrick et al., 2001; Hurtz & Donovan, 2000; Tett et al., 1991; Zell & Lesick, 2021), organizational citizenship behaviors (Chiaburu et al., 2011), counterproductive work behaviors (Salgado, 2002), and turnover (Zimmerman, 2008).

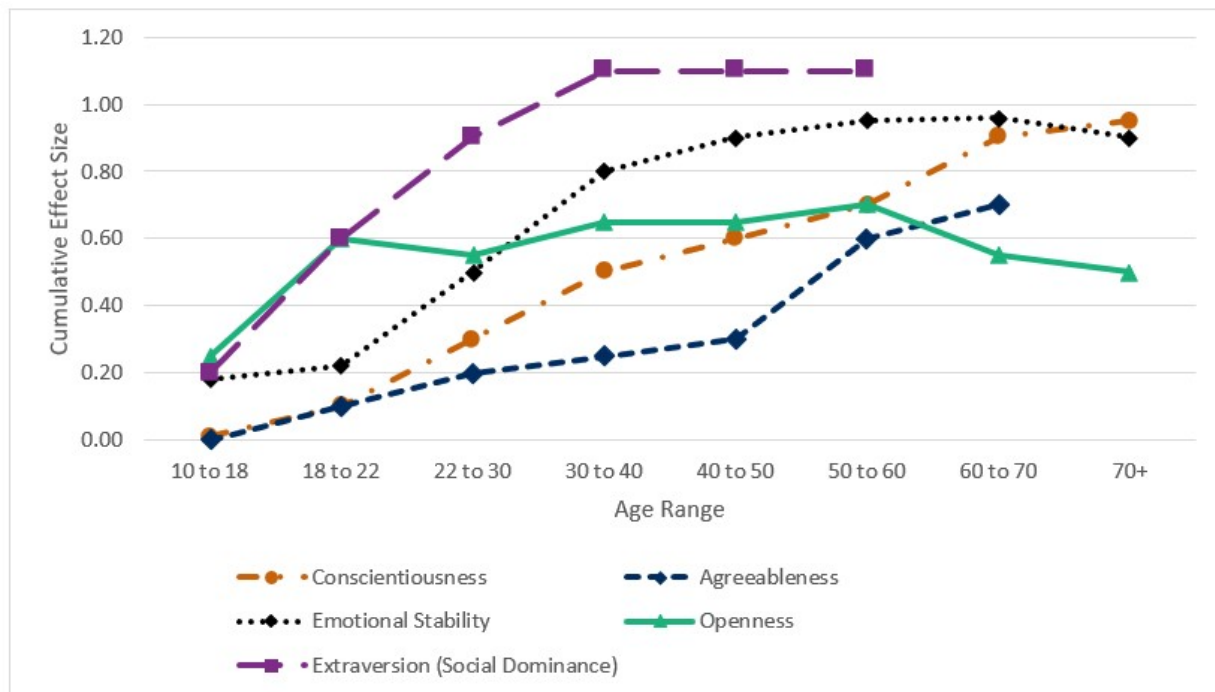
When it comes to task performance at work, research has shown that essential skills predict whether an individual completes tasks on time, how much work is completed, and what the quality and accuracy of that work are (e.g., Anderson et al., 1973; Thoresen et al., 2004; Viswesvaran et al., 1996). They also predict more personal outcomes, such as satisfaction, stress, and burnout, all of which can affect performance (e.g., Bogg & Roberts, 2004; Judge, Heller, & Mount, 2002; Swider & Zimmerman, 2010). Further, essential skills have been associated with a range of activities that contribute to the positive social and psychological aspects of an organization (Borman & Motowidlo, 1993) and with activities that generally involve more interpersonal tasks, such as working in teams, providing courteous service, managing conflict, and displaying leadership (e.g., Hogan & Holland, 2003; Judge, Bono, et al., 2002; Morgeson et al., 2005; Organ & Ryan, 1995; Sharma et al., 2013). Similarly, these skills have been negatively associated with inappropriate interpersonal conduct (often referred to as workplace incivility; Cortina et al., 2001), including discrimination, bullying, sexual harassment,

and other forms of aggression in the workplace (e.g., Bolton et al., 2010; Roberts, Harms, et al., 2007; Spector et al., 2006).

Five-Factor Model Traits' Development Over the Human Lifespan

Research supports the idea that these skills are important. This is especially true once an employee has been on the job for some time and learned all the major tasks required for that job (Murphy, 1989; Thoresen et al., 2004). At this point, what an employee *is willing to do*, which is in large part informed by essential skills such as work ethic, becomes more relevant to job performance than what an employee *can do*. However, little attention is paid to how these types of skills might be enhanced during an employee's career.

The fact that tremendous effort is not expended to develop employees' essential skills is likely due to the long-standing notion that personality is "set like plaster" (Costa & McCrae, 1994; James, 1981). That is, until recently, many believed that personality traits and essential skills were immutable, particularly as one aged. People's tendency to hold on to this view may stem from the fact that there is some trait consistency over time. For example, Roberts and DelVecchio (2000) examined 152 longitudinal studies to show that the rank-order consistency of personality was fairly moderate, with estimated test-retest correlations starting around .31 in early childhood (0–2.9 years) and generally increasing over time, reaching .74 in the 50–59 decade, when consistency estimates plateau. This level of consistency has been taken to indicate that personality does not change over time. However, high rank-order consistency does not imply that there is no mean-level change: Individuals can maintain their relative standing within a group over time while the entire group can shift a nontrivial degree. In their meta-analysis of mean-level personality change, Roberts et al. (2006) found that individuals become more socially dominant (a facet of extraversion), conscientious, agreeable, and emotionally stable throughout their lifespan, particularly in adolescence and early adulthood, and the effects were not slight; change over the lifespan reached a full standard deviation (see Figure 1).

Figure 1. Meta-Analytic Evidence Showing Personality Changes Over the Human Lifespan

Note. Roberts et al. (2006) examined the mean-level change of two components of extraversion with different developmental patterns, namely social dominance and social vitality. Here we present the findings for social dominance.

Research in this field has led to the conclusion that cognitive ability cannot be changed easily (Kyllonen et al., 2008), but personality traits can be, often fairly quickly and sometimes innocuously (Dweck, 2012). Roberts and colleagues carried out a meta-analysis showing that interventions (primarily clinical interventions) can alter personality traits to a significant degree even when the intervention lasts just eight weeks (Roberts et al., 2017). Outside the realm of clinical interventions, research suggests that personality trait change can result from something fairly simple: completing “challenges” (Hudson et al., 2019). The challenges Hudson and colleagues used were specific, concrete actions, small enough for someone to reasonably complete, such as “ask a friend to go for coffee” and “read a news story about a foreign country.” Completing two extraversion challenges per week, for example, resulted in an average increase of .17 standard deviations over the course of a single semester.

This body of research supports the potential for essential skill interventions. Although the effectiveness of essential skill development programs was once dubious, there is now a solid evidentiary base demonstrating that they are not only plausible but credible. Given the increasing value organizations are placing on essential skills, the effectiveness of high-quality essential skill development programs, and the relative cost effectiveness of such programs, essential skill development should play a more pivotal role in organizational development and effectiveness than it has previously.

WorkKeys Essential Skills and Their Alignment to the Five-Factor and HEXACO Models

As stated above, WorkKeys Essential Skills was designed to align with the widely accepted and supported Big Five and HEXACO taxonomies. Because of this, each skill assessed by WorkKeys Essential Skills can be directly linked to one of the Big Five or HEXACO constructs. Table 1 provides the WorkKeys Essential Skills constructs, their definitions, and their corresponding Big Five and HEXACO factors.

WorkKeys Essential Skills uses different naming conventions than those seen in the Big Five and HEXACO. This was done to better align the constructs to terminology commonly used within the workplace. These naming conventions will make the interpretation and communication of the assessment results more straightforward, and they will also help HR professionals, business leaders, and stakeholders understand how the constructs assessed by WorkKeys Essential Skills are directly relevant to important organizational criteria.

Table 1. Alignment of WorkKeys Essential Skills Domains to the Big Five and HEXACO Models

WorkKeys Essential Skills domain definition		
WorkKeys Essential Skills domain	<i>The extent to which a person's actions demonstrate . . .</i>	Big Five/HEXACO factor
Work ethic	Persistence, goal striving, reliability, dependability, and attention to detail at work	Conscientiousness
Collaboration	The ability to work on teams, empathy, helpfulness, trust, and trustworthiness	Agreeableness
Resilience	Stress management, emotional regulation, a positive response to setbacks, and poise	Emotional stability/emotionality
Creativity	Ingenuity, creative thinking, inquisitiveness, flexibility, open-mindedness, and embracing diversity	Openness to experience
Leadership	Assertiveness, influence, optimism, and enthusiasm	Extraversion
Integrity	Honesty, sincerity, fairness toward others, and modesty at work	Honesty-humility

ACT conducted this alignment by comparing the definitions of the WorkKeys Essential Skills constructs with those of the FFM factors. The alignment is supported by research demonstrating

significant correlations between the WorkKeys Essential Skills domains and the Big Five/HEXACO factors:

- Time management (Burrus et al., 2013) and goal setting (Klein & Lee, 2006), which are both related to work ethic, also correlate with conscientiousness.
- Interpersonal conflict (related to collaboration) correlates with agreeableness (Jensen-Campbell & Graziano, 2001).
- Test anxiety (Moutafi et al., 2006) and coping (MacCann et al., 2012), both of which are related to resilience, correlate with emotional stability.
- Curiosity (related to creativity) correlates with openness to experience (Komarraju et al., 2009), and creativity correlates with openness to experience (McCrae, 1987).
- Leadership correlates with extraversion (Judge, Bono, et al., 2002).

Some authors (e.g., Sackett & Walmsley, 2014) caution that the research literature (a) may not reflect what employers actually value and (b) contains convenience samples of occupations, which may not represent the full range of occupations in the workplace. As such, Sackett and Walmsley determined which essential skills are most valued by looking both at the types of questions job candidates are asked during structured interviews and at job analysis data collected by the Department of Labor's O*NET. The results of these analyses were consistent with those of meta-analyses. That is, when essential skills were ranked in terms of importance for a wide variety of job families, components from all six domains in the WorkKeys Essential Skills assessment were represented in the top three spots among all the job families (Sackett & Walmsley, 2014), with work ethic (conscientiousness) considered the most important. These results are also generally consistent with national surveys of employers, which have found that employers value essential skills in general and work ethic in particular (Casner-Lotto & Barrington, 2006; Cole et al., 2021).

Aligning WorkKeys Essential Skills to the ACT® Holistic Framework®

ACT has developed a holistic framework for education and workplace success; this framework provides a complete description of what people should know and be able and willing to do to succeed at school and in the workforce (Camara et al., 2015). The framework includes four broad domains:

1. *Core academic skills*: academic knowledge and skills from English language arts, mathematics, and science
2. *Cross-cutting capabilities*: knowledge and skills that bolster performance across academic areas (e.g., collaborative problem-solving)

3. *Behavioral skills*: interpersonal, self-regulatory, and task-related behaviors that are important for success in school and work
4. *Education and career navigation skills*: personal characteristics, knowledge, and processes that guide individuals through their education and career paths

Of these four domains, the behavioral skills domain is most relevant to WorkKeys Essential Skills. The behavioral skills framework is also based on the HEXACO model of personality (Ashton & Lee, 2007), and thus WorkKeys Essential Skills also aligns with the ACT Holistic Framework (Table 2).

Table 2. Alignment of WorkKeys Essential Skills to the ACT Behavioral Skills Framework

WorkKeys Essential Skills domain	ACT Behavioral Skills Framework domain (corresponding Big Five/HEXACO factor)
Work ethic	Sustaining effort (conscientiousness)
Collaboration	Getting along with others (agreeableness)
Resilience	Maintaining composure (emotional stability)
Creativity	Keeping an open mind (openness to experience)
Leadership	Socializing with others (extraversion)
Integrity	Acting honestly (honesty-humility)

Chapter 3: Multi-Trait Multi-Method Assessment

Several methods for measuring and scoring essential skills exist. Despite this, most assessments employ only one item type, most commonly the Likert item. Although Likert items have multiple advantages, they also have shortcomings that threaten their measurement accuracy (Kankaraš, 2017).

WorkKeys Essential Skills assesses the six essential skills with two methods: Likert items and situational judgment test (SJT) items. Every measurement method is subject to its own strengths and biases or weaknesses, and WorkKeys Essential Skills employs multiple methods to minimize the effects of the biases or weaknesses. This is known as a multi-trait multi-method (MTMM) design. According to Kenny and Kashy (1992), “the underlying view of measurement in the MTMM analysis is that to measure a theoretical construct, different measures, each with its own bias, are selected. Bias that is due to method effects is reduced through a triangulation process” (p. 170).

The benefit of the MTMM design is intuitive. For example, responses to Likert items can be easily faked (see below), whereas responses to SJTs are somewhat more difficult to fake. SJT items have problems of their own, however, such as some people having difficulty imagining themselves in hypothetical situations. If only one of these methods is employed in an assessment system, then the bias associated with that method can seriously compromise the assessment’s validity. The use of multiple methods minimizes this problem. Both assessment methods are briefly described below.

Likert Items

Likert items have been used in research and practice for decades and are capable of efficiently gathering a lot of information in a brief time. Individuals are asked to indicate their level of agreement with a number of statements.

However, respondents may have various motives for faking their responses, such as to avoid having to attend training programs or to appear more attractive to a prospective employer in a high-stakes selection context (e.g., Zickar et al., 2004). Furthermore, Likert items can be particularly susceptible to response biases. Reference bias, for example, involves people answering items by asking, “Compared to whom?” As a consequence, employees in high-skilled jobs might rate themselves lower on their essential skills than employees in low-skilled jobs simply because they are using a different reference group. This is often called the big-fish–little-pond effect (Marsh & Hau, 2003). Additional response biases include acquiescence and central tendency biases. With acquiescent responding, test takers tend to agree with all statements regardless of their content. In contrast, central tendency bias involves avoiding the ends of the response scale and focusing responses in the middle. All of these biases can affect the validity of an examinee’s score.

Situational Judgment Test Items

SJT items present a respondent with a situation and then several potential responses to that situation. Participants are asked how best to deal with those situations or how they would typically do so. The situations can be described in writing or depicted in audio/video, and response types include multiple choice, constructed response, and ratings (McDaniel et al., 2001). SJT items can measure a range of relevant attributes (e.g., leadership or teamwork) and have been shown to predict criteria such as leadership (Legree, 1995) and managerial performance (Howard & Choi, 2000).

SJTs have several advantages over Likert scales. First, while both Likert scales and SJTs can reflect general judgment processes, SJTs can reflect a greater number of subtle and complex processes than Likert scales can. Second, compared with cognitive assessments, SJTs appear to be associated with less adverse impact on ethnic minorities, which is relevant to mitigating subgroup differences in any population under consideration (Schmitt et al., 2009). Third, SJTs have the advantage of face validity; that is, the situations presented to employees and applicants look and feel like situations they would encounter in real life. Fourth, there is evidence suggesting SJTs are less prone to faking than Likert items (Hooper et al., 2006). Finally, respondents report that SJTs are engaging and worth completing (Lipnevich et al., 2013), which better supports multiple administrations and fosters employee investment in the ongoing process of essential skill assessment. SJTs are not without limitations, however. For example, they are often multidimensional (McDaniel & Whetzel, 2007) and may have a fairly high reading load, which can be taxing for the respondent.

Aggregate Score

The two item types are combined to yield one aggregate score per WorkKeys Essential Skills domain. To obtain the aggregate score for a domain, ACT first reverse-scores all negatively keyed items for that domain. The average of the six Likert items is then computed, as is the average of the six SJT items. Those mean scores are then standardized (*z*-scored), and the two are averaged to yield the aggregate score.

Norms

In reports provided to examinees, the aggregate scale scores are converted into three-level ratings for ease of interpretation. These are based on a normative sample with a mean age of 35.3 years (range = 18–78, *SD* = 11.06; see Table 3 for additional demographics). Level 1 corresponds to the bottom 16%. Level 2 corresponds to the 17th–84th percentiles. Level 3 corresponds to the 85th percentile and above. Those at Level 3 are eligible to receive virtual badges acknowledging their mastery of the domains.

Table 3. Normative Sample Demographics

Demographic Variable		Number of Participants
Gender	Female	822
	Male	1,156
	Other or Decline to answer	11
Race/Ethnicity	American Indian or Alaska Native	38
	Asian	431
	Black or African American	156
	Hispanic, Latino, or of Spanish origin	91
	Other	24
	White	1,249
Total	—	1,989

Do Not Report (DNR) Results

Respondents with inconsistent (high variance) and/or non-varied (low variance) response patterns are flagged. High variance likely indicates careless responding, while low variance likely indicates acquiescent responding or some other response bias. There are 14 possible response patterns that could be flagged across item types and domains (e.g., low variance across all Likert items, high variance across work ethic SJTs). If a respondent has two or more flags, they will not receive a score or report, as their response pattern is considered inconsistent.

Chapter 4: Initial WorkKeys Essential Skills Item Development

Initial Item Development

To generate the initial item pool, professional item writers were provided with WorkKeys Essential Skills construct definitions written by research scientists working on the development of the assessment. Items were contextualized so that the scenarios would fit an employee's experience. The research scientists reviewed each item and provided feedback, and the items were revised accordingly. Additional item revisions or selections were made based on data collected from 1,768 adults who completed a pilot version of WorkKeys Essential Skills. Items were removed if they exhibited poor psychometric qualities (e.g., they failed to load highly on their target factor, or they had poor validity estimates).

Final Item Set

The final WorkKeys Essential Skills item set includes 36 Likert items (six per construct) and 12 SJTs (two per construct, each with three behavioral responses). The set has a Flesch-Kincaid reading level of 5.4. Table 4 provides an example of each item type.

Table 4. WorkKeys Essential Skills Item Types

Item type	Example item	Number of items
Likert	Indicate the extent to which you agree or disagree with the statement on a scale from 1 (strongly disagree) to 6 (strongly agree). <i>I cooperate with others.</i>	Total: 36 • Six items per domain
Situational judgment test	Read the situation described, then rate how likely you are to do each of the actions that follow on a scale from 1 (very unlikely) to 5 (very likely). <i>You have been assigned to train recently hired coworkers. During the first day of training, you realize that some of the new coworkers are grasping the material quickly, but a few are really struggling.</i> <ul style="list-style-type: none"> • <i>Tell the struggling trainees that they need to work harder.</i> • <i>Praise the best trainees in front of those who are struggling to motivate them to catch up.</i> • <i>Tell the struggling trainees that they may want to apply for a new position.</i> 	Total: 12 situations, each with three behavioral responses • Two situations per domain, each with three behavioral responses

Chapter 5: WorkKeys Essential Skills Reliability and Validity Evidence

In this chapter, we provide the results of a content validity investigation, as well as findings from two pilot studies supporting the use of the WorkKeys Essential Skills assessment.

Content Validity Evidence

Validity is defined as “the degree to which evidence and theory support the interpretations of test scores for proposed uses of tests” (*Standards for Educational and Psychological Testing*; American Educational Research Association [AERA] et al., 2014, p. 11). Validity can be thought of as a unitary concept with multiple sources of evidence that can support claims about how test scores should be interpreted given a test’s proposed uses. When there is strong validity evidence, we can be more confident that we are drawing the correct conclusions from an assessment. To validate WorkKeys Essential Skills, ACT researchers considered all evidence of content, construct, and criterion validity (Society for Industrial and Organizational Psychology [SIOP], 2018). Content validity evidence is discussed below, construct validity evidence is discussed in Pilot Study #1, and criterion validity evidence is discussed in Pilot Study #2.

Content validity concerns the relationship between a test’s content and the constructs that the test intends to measure. Evaluations of content validity can include both logical and empirical analyses of how well the test items map to the constructs they are intended to measure. WorkKeys Essential Skills was designed to align with the FFM constructs and the honesty-humility construct of the HEXACO model. As mentioned in Chapter 2, subject matter experts (SMEs) based the WorkKeys Essential Skills construct descriptions on the Big Five and HEXACO descriptions to establish crosswalks between the two frameworks (shown in Table 1). Item writers then wrote assessment items based on these skill descriptions. The content assessed by WorkKeys Essential Skills should therefore be representative of the content captured by the FFM and the HEXACO honesty-humility construct.

In order to examine this rigorously, six SMEs with doctoral degrees in psychology and extensive knowledge of the FFM and HEXACO models independently examined the content of each WorkKeys Essential Skills item and identified the FFM (or honesty-humility) factor whose content best fit that item. All six SMEs identified the intended factor for 31 out of 36 Likert items, five identified the factor for four items, and four identified the factor for the final item. All six SMEs correctly identified the factor for 11 out of 12 SJTs, and four correctly identified it for the final SJT. When all of the SME ratings for the items were taken together by type, Likert items were judged to best fit the content of the intended factor in 97% of cases, and SJT items were also judged to best fit the content of the intended factor in 97% of cases. Therefore, the content of the WorkKeys Essential Skills scales does indeed represent the content captured by the FFM and the HEXACO honesty-humility construct.

Pilot Study #1: Reliability and Construct Validity Evidence

Overview

Below we report reliability and validity evidence collected from adults who participated in an online study. Participants from the sample responded to the WorkKeys Essential Skills assessment and to two validated, well-established measures of the Big Five and HEXACO frameworks: the 10-item Big Five Inventory (BFI-10; Rammstedt & John, 2007) and the honesty-humility scale from the HEXACO-PI (HEXACO-H; Lee & Ashton, 2018). This enabled us to evaluate the assessment's construct validity evidence.

Method

Participants

The study was conducted using an online Amazon Mechanical Turk (MTurk) sample of 1,768 adults with a mean age of 36 years ($SD = 11.1$). Demographic information can be found in Table 5.

Table 5. Pilot Study #1 Demographics

Demographic Variable		Number of Participants
Gender	Female	744
	Male	1,014
	Other or Decline to answer	10
Race/ethnicity	American Indian or Alaska Native	27
	Asian	359
	Black or African American	134
	Hispanic, Latino, or of Spanish origin	83
	Other	19
	White	1,146
Total	—	1,768

Results

Evidence for Reliability

The reliability (or precision) of an assessment refers to the consistency of scores obtained from the assessment (AERA et al., 2014; SIOP, 2018). Reliability evidence is always important to consider in test development, and the need for precision increases when high-stakes decisions are made based on test scores.

Internal Consistency Reliability

The internal consistency estimates of the Likert scales were as follows: work ethic = .85, collaboration = .82, resilience = .89, creativity = .87, leadership = .88, and integrity = .77. The

internal consistency estimates of the SJT scales were as follows: work ethic = .71, collaboration = .87, resilience = .82, creativity = .74, leadership = .79, and integrity = .82. The Likert and SJT scale scores were standardized, and Cronbach's alpha was estimated for the aggregate scores. The scales' reliability estimates were as follows: work ethic = .66, collaboration = .32, resilience = .79, creativity = .74, leadership = .60, and integrity = .35.

Evidence for Validity

Construct Validity Evidence

Construct validity is the extent to which a test correlates with similar and dissimilar measures, and it can be evaluated through examining convergent and discriminant validity evidence. AERA et al. state that "the patterns of association between and among scores on the test under study and other variables should be consistent with theoretical expectations" (2014, p. 27). Because each WorkKeys Essential Skills construct was designed to theoretically align with its corresponding Big Five or HEXACO factor, each essential skill should correlate most highly with the trait it was designed to align with (convergent validity) while correlating to a lesser degree with other traits measured by the BFI-10 and HEXACO-H (discriminant validity). Thus, correlations among the group of scale scores (six constructs of WorkKeys Essential Skills by five constructs of the BFI-10 plus HEXACO-H) were computed to evaluate evidence for convergent and discriminant validity.

Each WorkKeys Essential Skills construct correlated most highly with its corresponding trait on the BFI-10 and HEXACO-H and had smaller correlations with the other BFI and HEXACO-H traits (see Table 6). For example, the WorkKeys Essential Skills resilience skill (designed to measure emotional stability) correlated with BFI-10 emotional stability at .80 but correlated with all other traits at .42 and under. There was one exception: Integrity correlated .54 with honesty-humility but .53 with conscientiousness and .52 with agreeableness. However, this makes sense because in the Big Five model, integrity falls under these two traits. Overall, this analysis provides evidence of convergent and discriminant validity for WorkKeys Essential Skills. By establishing construct-related validity evidence, we increase our confidence that the assessment is accurately measuring the appropriate skills.

Table 6. Correlations Between WorkKeys Essential Skills Domains and BFI-10/HEXACO-H

WorkKeys Essential Skills domain	Big Five Inventory or HEXACO-H factor					
	Conscientiousness	Agreeableness	Emotional stability	Openness to experience	Extraversion	Honesty-humility
Work ethic	.65*	.44	.33	.34	.10	.37
Collaboration	.47	.56*	.29	.40	.04	.49
Resilience	.42	.33	.80*	.33	.32	.17
Creativity	.23	.22	.28	.50*	.37	-.01
Leadership	.26	.20	.44	.38	.69*	-.10
Integrity	.53	.52	.35	.37	.05	.54*

* Correlation between the WorkKeys Essential Skills domain and its respective BFI-10 or HEXACO-H factor

Subgroup Differences

ACT researchers evaluated demographic group differences to determine whether there were any significant subgroup differences on the WorkKeys Essential Skills scales. The researchers conducted mean comparisons using independent samples *t*-tests. Comparisons were made between gender groups, racial/ethnic groups, and age groups. The results of these analyses are shown in Table 7 in terms of the standardized effect size of the mean differences between groups for each comparison.

Gender Differences

Comparing male to female respondents showed significant differences on five of the six WorkKeys Essential Skills constructs. Female respondents scored higher on work ethic, collaboration, and integrity, whereas male respondents scored higher on resilience and leadership.

Race/Ethnicity Differences

Comparing White to Black respondents showed significant differences on creativity and leadership, with Black respondents scoring higher than White respondents. Comparing White to Hispanic respondents showed no significant differences on any of the skills. Comparing White to Asian respondents showed large, significant differences on five of the six skills. White respondents scored higher on work ethic, collaboration, and integrity, whereas Asian respondents scored higher on creativity and leadership.

Age Differences

Comparing respondents younger than 40 to those 40 or older (40+) showed significant differences on four of the six skills. The 40+ respondents scored higher on work ethic, collaboration, resilience, and integrity.

Table 7. Subgroup Differences on WorkKeys Essential Skills Domains

WorkKeys Essential Skills domain	Subgroup comparisons				
	Male–female	White–Black	White–Hispanic	White–Asian	Below 40 years–40+ years
Work ethic	–.38*	.06	.02	.50*	–.38*
Collaboration	–.56*	.06	.14	.86*	–.35*
Resilience	.29*	–.09	–.14	.03	–.24*
Creativity	.08	–.30*	–.09	–.21*	.06
Leadership	.29*	–.28*	–.05	–.31*	–.00
Integrity	–.48*	.02	.12	.69*	–.45*

Note. Values are Cohen’s *d* values for effect size of mean differences. * indicates a significant mean difference at $p < .05$. A positive value means the male, White, or below 40 group is higher than the historically underrepresented counterpart on the skill.

Pilot Study #2: Criterion-Related Validity Evidence

Overview

The second WorkKeys Essential Skills pilot study collected evidence of the assessment's ability to predict job performance. The job performance measure in this study was developed by ACT for criterion-validity research with its workforce line of assessments (ACT, 2015).

Method

Participants

The study sample included 284 participants with a mean age of 43 years ($SD = 13.2$). Demographic information for all participants can be found in Table 8. Correlations between a 24-item measure of supervisor-rated job performance (Cronbach's alpha = .96) and the six WorkKeys Essential Skills scale scores were examined in a subset of 173 employees who had matched supervisor performance ratings.

Table 8. Pilot Study #2 Demographics

Demographic variable		Number of employees
Gender	Female	195
	Male	87
	Other or Decline to answer	2
Race/ethnicity	American Indian or Alaska Native	2
	Asian	6
	Black or African American	55
	Hispanic, Latino, or of Spanish origin	17
	Other	12
	White	192
Hours worked per week	Fewer than 40 hours per week	36
	40 or more hours per week	248
Occupation	Education	17
	Government agency	69
	Healthcare	1
	Manufacturing	44
	Professional	31
	Retail or service	6
	Technical	2
	Other	114
Total	—	284

Results

Evidence for Validity

One method for establishing the validity of a measure is to collect evidence that supports its job-relatedness (SIOP, 2018). For an assessment to be used in high-stakes selection contexts, there must be evidence that applicant scores on the assessment are related to the individual and/or organizational outcomes that the assessment is intended to predict (typically job performance). Establishing criterion-related validity involves examining evidence regarding the relationship between the predictor (WorkKeys Essential Skills) and the criterion (job performance). Criterion measures are separate constructs that are hypothesized to be related to constructs measured by the test in question. A statistical relationship between WorkKeys Essential Skills and the job-performance criterion measure would be evidence of criterion-related validity (SIOP, 2018).

Criterion-Related Validity Evidence

In line with meta-analytic findings (Barrick et al., 2001; Zell & Lesick, 2021), work ethic (conscientiousness) had the strongest relationship with job performance ($r = .29$). Collaboration (agreeableness; $r = .22$) also had a statistically significant relationship with job performance, as did creativity (openness to experience; $r = .17$). One main difference between the current results and previous meta-analytic findings is the lack of significance for the resilience scale ($r = .09$). Leadership ($r = .11$) and integrity ($r = .03$) also had small, nonsignificant correlations with job performance. Other scales may also be significant predictors of performance in individual organizations depending on the job (e.g., sales) or criterion (e.g., training performance).

References

- ACT. (2015). *Summary of WorkKeys criterion validity research* [Unpublished report]. ACT.
- ACT. (2021). *Mosaic™ by ACT®: Social Emotional Learning Assessment*.
- Allport, G. W., & Odbert, H. S. (1936). Trait-names: A psycho-lexical study. *Psychological Monographs*, 47(1), i–171. <https://doi.org/10.1037/h0093360>
- American Educational Research Association, American Psychological Association, & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. AERA.
- Anderson, H. E., Roush, S. L., & McClary, J. E. (1973). Relationships among ratings, production, efficiency, and the General Aptitude Test Battery scales in an industrial setting. *Journal of Applied Psychology*, 58(1), 77–82. <https://doi.org/10.1037/h0035413>
- Ashton, M. C., & Lee, K. (2007). Empirical, theoretical, and practical advantages of the HEXACO model of personality structure. *Personality and Social Psychology Review*, 11(2), 150–166. <https://doi.org/10.1177/1088868306294907>
- Ashton, M. C., Lee, K., Perugini, M., Szarota, P., de Vries, R. E., di Blas, L., Boies, K., & de Raad, B. (2004). A six-factor structure of personality-descriptive adjectives: Solutions from psycholexical studies in seven languages. *Journal of Personality and Social Psychology*, 86(2), 356–366. <https://doi.org/10.1037/0022-3514.86.2.356>
- Barrick, M. R., & Mount, M. K. (1991). The Big Five personality dimensions and job performance: A meta-analysis. *Personnel Psychology*, 44(1), 1–26. <https://doi.org/10.1111/j.1744-6570.1991.tb00688.x>
- Barrick, M. R., Mount, M. K., & Judge, T. A. (2001). Personality and performance at the beginning of the new millennium: What do we know and where do we go next? *International Journal of Selection and Assessment*, 9(1/2), 9–30. <https://doi.org/10.1111/1468-2389.00160>
- Bogg, T., & Roberts, B. W. (2004). Conscientiousness and health-related behaviors: A meta-analysis of the leading behavioral contributors to mortality. *Psychological Bulletin*, 130(6), 887–919. <https://doi.org/10.1037/0033-2909.130.6.887>

- Bolton, L. R., Becker, L. K., & Barber, L. K. (2010). Big Five trait predictors of differential counterproductive work behavior dimensions. *Personality and Individual Differences*, 49(5), 537–541. <https://doi.org/10.1016/j.paid.2010.03.047>
- Borman, W. C., & Motowidlo, S. J. (1993). Expanding the criterion domain to include elements of contextual performance. In N. Schmitt & W. C. Borman (Eds.), *Personnel selection in organizations* (pp. 71–98). Jossey-Bass.
- Bughin, J., Hazan, E., Lund, S., Dahlström, P., Wiesinger, A., & Subramaniam, A. (2018, May 23). *Skill shift: Automation and the future of the workforce*. McKinsey & Company. <https://www.mckinsey.com/featured-insights/future-of-work/skill-shift-automation-and-the-future-of-the-workforce>
- Burrus, J., Jackson, T., Holtzman, S., Roberts, R. D., & Mandigo, T. (2013). *Examining the efficacy of a time management intervention for high school students* [ETS Research Report 13-25]. Educational Testing Service. <http://dx.doi.org/10.1002/j.2333-8504.2013.tb02332.x>
- 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 2015-4]. ACT. https://www.act.org/content/dam/act/unsecured/documents/ACT_RR2015-4.pdf
- Casner-Lotto, J., & Barrington, L. (2006). *Are they really ready to work?: Employers' perspectives on the basic knowledge and applied skills of new entrants to the 21st century U.S. workforce*. The Conference Board; Partnership for 21st Century Skills; Corporate Voices for Working Families; Society for Human Resource Management. <https://files.eric.ed.gov/fulltext/ED519465.pdf>
- Cattell, R. B. (1943). The description of personality: The foundations of trait measurement. *Psychological Review*, 50(6), 559–594. <https://doi.org/10.1037/h0057276>
- Chiaburu, D. S., Oh, I. S., Berry, C. M., Li, N., & Gardner, R. G. (2011). The five-factor model of personality traits and organizational citizenship behaviors: A meta-analysis. *Journal of Applied Psychology*, 96(6), 1140–1166. <https://doi.org/10.1037/a0024004>

- Cole, L., Short, S., Cowart, C., & Muller, S. (2021). *The high demand for durable skills*. America Succeeds. <https://americasucceeds.org/wp-content/uploads/2021/04/AmericaSucceeds-DurableSkills-NationalFactSheet-2021.pdf>
- Cortina, L. M., Magley, V. J., Williams, J. H., & Langout, R. D. (2001). Incivility in the workplace: Incidence and impact. *Journal of Occupational Health Psychology, 6*(1), 64–80. <https://doi.org/10.1037/1076-8998.6.1.64>
- Costa, P. T., Jr., & McCrae, R. R. (1994). Set like plaster? Evidence for the stability of adult personality. In T. F. Heatherton & J. L. Weinberger (Eds.), *Can personality change?* (pp. 21–40). American Psychological Association. <https://psycnet.apa.org/doi/10.1037/10143-002>
- Cronbach, L. J., & Gleser, G. C. (1965). *Psychological tests and personnel decisions* (2nd ed.). University of Illinois Press.
- de Raad, B., & Mlačić, B. (2015). The lexical foundation of the Big Five Factor Model. In T. A. Widiger (Ed.), *The Oxford handbook of the Five Factor Model* (pp. 191–216). Oxford University Press.
- Digman, J. M. (1990). Personality structure: Emergence of the five-factor model. *Annual Review of Psychology, 41*, 417–440. <https://doi.org/10.1146/annurev.ps.41.020190.002221>
- Dweck, C. S. (2012). Mindsets and human nature: Promoting change in the Middle East, the schoolyard, the racial divide, and willpower. *American Psychologist, 67*(8), 614–622. <https://doi.org/10.1037/a0029783>
- Fiske, D. W. (1949). Consistency of the factorial structures of personality ratings from different sources. *Journal of Abnormal and Social Psychology, 44*(3), 329–344. <https://doi.org/10.1037/h0057198>
- Goldberg, L. R. (1993). The structure of phenotypic personality traits. *American Psychologist, 48*(1), 26–34. <https://doi.org/10.1037/0003-066X.48.1.26>
- Hogan, J., & Holland, B. (2003). Using theory to evaluate personality and job-performance relations: A socioanalytic perspective. *Journal of Applied Psychology, 88*(1), 100–112. <https://doi.org/10.1037/0021-9010.88.1.100>
- Hooper, A. C., Cullen, M. J., & Sackett, P. R. (2006). Operational threats to the use of SJTs: Faking, coaching, and retesting issues. In J. A. Weekley & R. E. Ployhart

- (Eds.), *Situational judgment tests: Theory, measurement, and application* (pp. 205–232). Psychology Press.
- Howard, A., & Choi, M. (2000). How do you assess a manager's decision-making abilities? The use of situational inventories. *International Journal of Selection and Assessment*, 8(2), 85–88. <https://doi.org/10.1111/1468-2389.00136>
- Hudson, N. W., Briley, D. A., Chopik, W. J., & Derringer, J. (2019). You have to follow through: Attaining behavioral change goals predicts volitional personality change. *Journal of Personality and Social Psychology*, 117(4), 839-857. <https://doi.org/10.1037/pspp0000221>
- Hurtz, G. M., & Donovan, J. J. (2000). Personality and job performance: The Big Five revisited. *Journal of Applied Psychology*, 85(6), 869–879. <https://doi.org/10.1037/0021-9010.85.6.869>
- James, W. (1981). *The principles of psychology* (Vol. 1). Harvard University Press.
- Jensen-Campbell, L. A., & Graziano, W. G. (2001). Agreeableness as a moderator of interpersonal conflict. *Journal of Personality*, 69(2), 323–362. <https://doi.org/10.1111/1467-6494.00148>
- John, O. P., Naumann, L. P., & Soto, C. J. (2008). Paradigm shift to the integrative Big Five trait taxonomy: History, measurement, and conceptual issues In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality: Theory and research* (pp. 114–158). The Guilford Press.
- Johnson, M. K., Rowatt, W. C., & Petrini, L. (2011). A new trait on the market: Honesty–humility as a unique predictor of job performance ratings. *Personality and Individual Differences*, 50(6), 857–862. <https://doi.org/10.1016/j.paid.2011.01.011>
- Judge, T. A., Bono, J. E., Ilies, R., & Gerhardt, M. W. (2002). Personality and leadership: A qualitative and quantitative review. *Journal of Applied Psychology*, 87(4), 765–780. <https://doi.org/10.1037/0021-9010.87.4.765>
- Judge, T. A., Heller, D., & Mount, M. K. (2002). Five-factor model of personality and job satisfaction: A meta-analysis. *Journal of Applied Psychology*, 87(3), 530–541. <https://doi.org/10.1037/0021-9010.87.3.530>

- Judge, T. A., Higgins, C., Thoresen, C. J., & Barrick, M. R. (1999). The Big Five personality traits, general mental ability, and career success across the life span. *Personnel Psychology*, 52(3), 621–652. <https://doi.org/10.1111/j.1744-6570.1999.tb00174.x>
- Kankaraš, M. (2017). *Personality matters*. Organisation for Economic Co-operation and Development.
- Kenny, D. A., & Kashy, D. A. (1992). Analysis of the multitrait–multimethod matrix by confirmatory factor analysis. *Psychological Bulletin*, 112(1), 165–172. <https://doi.org/10.1037/0033-2909.112.1.165>
- Klein, H. J., & Lee, S. (2006). The effects of personality on learning: The mediating role of goal setting. *Human Performance*, 19(1), 43–66. https://doi.org/10.1207/s15327043hup1901_3
- Komarraju, M., Karau, S. J., & Schmeck, R. R. (2009). Role of the Big Five personality traits in predicting college students' academic motivation and achievement. *Learning and Individual Differences*, 19(1), 47–52. <https://doi.org/10.1016/j.lindif.2008.07.001>
- Kyllonen, P. C., Lipnevich, A. A., Burrus, J., & Roberts, R. D. (2014). *Personality, motivation, and college readiness: A prospectus for assessment and development* [ETS Research Report 14-06]. Educational Testing Service. <https://files.eric.ed.gov/fulltext/EJ1109311.pdf>
- Kyllonen, P. C., Roberts, R. D., & Stankov, L. (Eds.). (2008). *Extending intelligence: Enhancement and new constructs*. Routledge.
- Lee, K., & Ashton, M. C. (2018). Psychometric properties of the HEXACO-100. *Assessment*, 25(5), 543–556. <https://doi.org/10.1177/1073191116659134>
- Lee, K., Ashton, M. C., & de Vries, R. E. (2005). Predicting workplace delinquency and integrity with the HEXACO and Five-Factor models of personality structure. *Human Performance*, 18, 179–197. https://doi.org/10.1207/s15327043hup1802_4
- Lee, K., Ashton, M. C., Pozzebon, J. A., Visser, B. A., Bourdage, J. S., & Ogunfowora, B. (2009). Similarity and assumed similarity in personality reports of well-acquainted persons. *Journal of Personality and Social Psychology*, 96(2), 460–472. <https://doi.org/10.1037/a0014059>

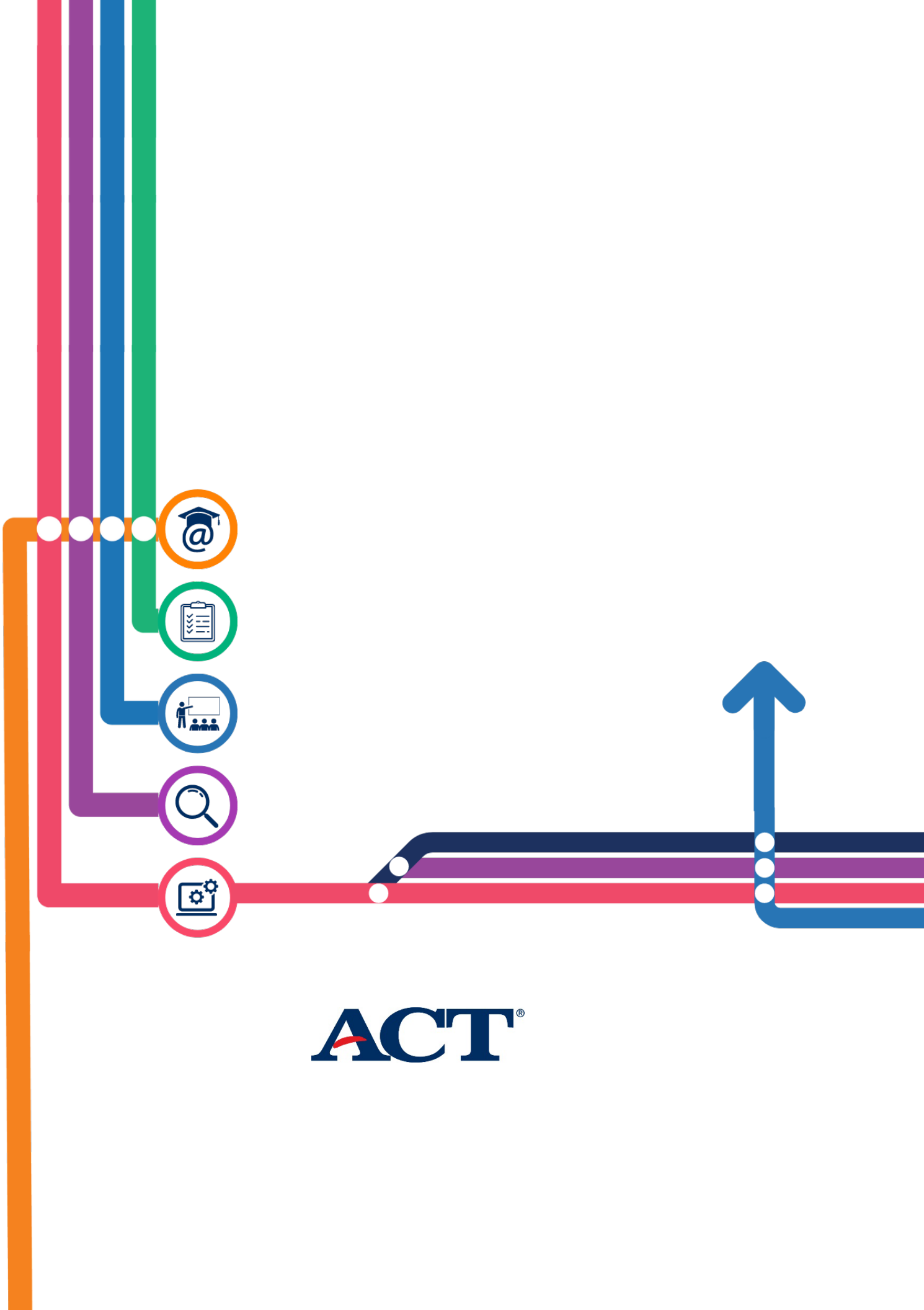
- Legree, P. J. (1995). Evidence for an oblique social intelligence factor established with a Likert-based testing procedure. *Intelligence*, 21(3), 247–266. [https://doi.org/10.1016/0160-2896\(95\)90016-0](https://doi.org/10.1016/0160-2896(95)90016-0)
- Lindqvist, E., & Vestman, R. (2011). The labor market returns to cognitive and noncognitive ability: Evidence from the Swedish enlistment. *American Economic Journal: Applied Economics*, 3(1), 101–128. <https://doi.org/10.1257/app.3.1.101>
- Lipnevich, A. A., MacCann, C., & Roberts, R. D. (2013). Assessing noncognitive constructs in education: A review of traditional and innovative approaches. In D. H. Saklofske, C. R. Reynolds, & V. L. Schwann (Eds.), *The Oxford handbook of child psychological assessment* (pp. 750–772). Oxford University Press.
- MacCann, C., Lipnevich, A. A., Burrus, J., & Roberts, R. D. (2012). The best years of our lives? Coping with stress predicts school grades, life satisfaction, and feelings about high school. *Learning and Individual Differences*, 22(2), 235–241. <https://doi.org/10.1016/j.lindif.2011.08.004>
- Mammadov, S. (2021). Big Five personality traits and academic performance: A meta-analysis. *Journal of Personality*, 90, 222–255. <https://doi.org/10.1111/jopy.12663>
- Marsh, H. W., & Hau, K. T. (2003). Big-fish–little-pond effect on academic self-concept: A cross-cultural (26-country) test of the negative effects of academically selective schools. *American Psychologist*, 58(5), 364–376. <https://doi.org/10.1037/0003-066X.58.5.364>
- McCrae, R. R. (1987). Creativity, divergent thinking, and openness to experience. *Journal of Personality and Social Psychology*, 52(6), 1258–1265. <https://doi.org/10.1037/0022-3514.52.6.1258>
- McCrae, R. R., Terracciano, A., & 78 Members of the Personality Profiles of Cultures Project. (2005). Universal features of personality traits from the observer’s perspective: Data from 50 cultures. *Journal of Personality and Social Psychology*, 88(3), 547–561. <https://doi.org/10.1037/0022-3514.88.3.547>
- McDaniel, M. A., Morgeson, F. P., Finnegan, E. B., Campion, M. A., & Braverman, E. P. (2001). Use of situational judgment tests to predict job performance: A clarification of the

- literature. *Journal of Applied Psychology*, 86(4), 730–740. <https://doi.org/10.1037/0021-9010.86.4.730>
- McDaniel, M. A., & Whetzel, D. L. (2007). Situational judgment tests. In D. L. Whetzel & G. R. Wheaton (Eds.), *Applied measurement: Industrial psychology in human resources management* (pp. 235–257). Taylor & Francis Group/Lawrence Erlbaum Associates.
- Morgeson, F. P., Reider, M. H., & Campion, M. A. (2005). Selecting individuals in team settings: The importance of social skills, personality characteristics, and teamwork knowledge. *Personnel Psychology*, 58(3), 583–611. <https://doi.org/10.1111/j.1744-6570.2005.655.x>
- Moutafi, J., Furnham, A., & Tsaousis, I. (2006). Is the relationship between intelligence and trait neuroticism mediated by test anxiety? *Personality and Individual Differences*, 40(3), 587–597. <https://doi.org/10.1016/j.paid.2005.08.004>
- Murphy, K. R. (1989). Is the relationship between cognitive ability and job performance stable over time? *Human Performance*, 2(3), 183–200. https://doi.org/10.1207/s15327043hup0203_3
- Norman, W. T. (1963). Toward an adequate taxonomy of personality attributes: Replicated factor structure in peer nomination personality ratings. *Journal of Abnormal and Social Psychology*, 66(6), 574–583. <https://doi.org/10.1037/h0040291>
- Ones, D. S., Viswesvaran, C., & Schmidt, F. L. (1993). Comprehensive meta-analysis of integrity test validities: Findings and implications for personnel selection and theories of job performance. *Journal of Applied Psychology*, 78(4), 679–703. <https://doi.org/10.1037/0021-9010.78.4.679>
- Organ, D. W., & Ryan, K. (1995). A meta-analytic review of attitudinal and dispositional predictors of organizational citizenship behavior. *Personnel Psychology*, 48(4), 775–802. <https://doi.org/10.1111/j.1744-6570.1995.tb01781.x>
- Poropat, A. E. (2009). A meta-analysis of the five-factor model of personality and academic performance. *Psychological Bulletin*, 135(2), 322–338. <https://doi.org/10.1037/a0014996>
- Rammstedt, B., & John, O. P. (2007). Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. *Journal of Research in Personality*, 41, 203–212. <https://doi.org/10.1016/j.jrp.2006.02.001>

- Roberts, B. W., & DeVecchio, W. F. (2000). The rank-order consistency of personality traits from childhood to old age: A quantitative review of longitudinal studies. *Psychological Bulletin*, 126(1), 3–25. <https://doi.org/10.1037/0033-2909.126.1.3>
- Roberts, B. W., Harms, P. D., Caspi, A., & Moffitt, T. E. (2007). Predicting the counterproductive employee in a child-to-adult prospective study. *Journal of Applied Psychology*, 92(5), 1427–1436. <https://doi.org/10.1037/0021-9010.92.5.1427>
- Roberts, B. W., Kuncel, N. R., Shiner, R., Caspi, A., & Goldberg, L. R. (2007). The power of personality: The comparative validity of personality traits, socioeconomic status, and cognitive ability for predicting important life outcomes. *Perspectives on Psychological Science*, 2(4), 313–345. <https://doi.org/10.1111/j.1745-6916.2007.00047.x>
- Roberts, B. W., Luo, J., Briley, D. A., Chow, P. I., Su, R., & Hill, P. L. (2017). A systematic review of personality trait change through intervention. *Psychological Bulletin*, 143(2), 117–141. <https://doi.org/10.1037/bul0000088>
- Roberts, R. D., Martin, J. E., & Olaru, G. (2015). *A Rosetta stone for noncognitive skills: Understanding, assessing, and enhancing noncognitive skills in primary and secondary education*. Asia Society; Professional Examination Service.
https://asiasociety.org/files/A_Rosetta_Stone_for_Noncognitive_Skills.pdf
- Roberts, B. W., Walton, K. E., & Viechtbauer, W. (2006). Patterns of mean-level change in personality traits across the life course: a meta-analysis of longitudinal studies. *Psychological Bulletin*, 132(1), 1–25. <https://doi.org/10.1037/0033-2909.132.1.1>
- Sackett, P. R., & Walmsley, P. T. (2014). Which personality attributes are most important in the workplace? *Perspectives on Psychological Science*, 9(5), 538–551.
<https://doi.org/10.1177/1745691614543972>
- Salgado, J. F. (2002). The Big Five personality dimensions and counterproductive behaviors. *International Journal of Selection and Assessment*, 10(1-2), 117–125.
<https://doi.org/10.1111/1468-2389.00198>
- Sharma, S., Bottom, W. P., & Elfenbein, H. A. (2013). On the role of personality, cognitive ability, and emotional intelligence in predicting negotiation outcomes: A meta-analysis. *Organizational Psychology Review*, 3(4), 293–336.
<https://doi.org/10.1177/2041386613505857>

- Schmidt, F. L., & Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin*, *124*(2), 262–274. <https://doi.org/10.1037/0033-2909.124.2.262>
- Schmitt, D., Allik, J., McCrae, R. R., & Benet-Martinez, V. (2007). The geographic distribution of Big Five personality traits: Patterns and profiles of human self-description across 56 nations. *Journal of Cross-Cultural Psychology*, *38*(2), 173–212. <https://doi.org/10.1177/0022022106297299>
- Schmitt, N., Keeney, J., Oswald, F. L., Pleskac, T. J., Billington, A. Q., Sinha, R., & Zorzie, M. (2009). Prediction of 4-year college student performance using cognitive and noncognitive predictors and the impact on demographic status of admitted students. *Journal of Applied Psychology*, *94*(6), 1479–1497. <https://doi.org/10.1037/a0016810>
- Society for Industrial and Organizational Psychology. (2018). *Principles for the validation and use of personnel selection procedures* (5th ed.). APA.
- Soto, C. J., & John, O. P. (2014). Traits in transition: The structure of parent-reported personality traits from early childhood to early adulthood. *Journal of Personality*, *82*(3), 182–199. <https://doi.org/10.1111/jopy.12044>
- Spector, P. E., Fox, S., Penney, L. M., Bruursema, K., Goh, A., & Kessler, S. (2006). The dimensionality of counterproductivity: Are all counterproductive behaviors created equal? *Journal of Vocational Behavior*, *68*(3), 446–460. <https://doi.org/10.1016/j.jvb.2005.10.005>
- Swider, B. W., & Zimmerman, R. D. (2010). Born to burnout: A meta-analytic path model of personality, job burnout, and work outcomes. *Journal of Vocational Behavior*, *76*(3), 487–506. <https://doi.org/10.1016/j.jvb.2010.01.003>
- Tett, R. P., Jackson, D. N., & Rothstein, M. (1991). Personality measures as predictors of job performance: A meta-analytic review. *Personnel Psychology*, *44*(4), 703–742. <https://doi.org/10.1111/j.1744-6570.1991.tb00696.x>
- Thoresen, C. J., Bradley, J. C., Bliese, P. D., & Thoresen, J. D. (2004). The Big Five personality traits and individual job performance growth trajectories in maintenance and transitional

- job stages. *Journal of Applied Psychology*, 89(5), 835–853. <https://doi.org/10.1037/0021-9010.89.5.835>
- Tupes, E. C., & Christal, R. C. (1961). *Recurrent personality factors based on trait ratings* (USAF ASD Technical Report 61-97). U.S. Air Force. <https://apps.dtic.mil/sti/tr/pdf/AD0267778.pdf>
- van Iddekinge, C. H., Roth, P. L., Raymark, P. H., & Odle-Dusseau, H. N. (2012). The criterion-related validity of integrity tests: An updated meta-analysis. *Journal of Applied Psychology*, 97(3), 499–530. <https://doi.org/10.1037/a0021196>
- Viswesvaran, C., Ones, D. S., & Schmidt, F. L. (1996). Comparative analysis of the reliability of job performance ratings. *Journal of Applied Psychology*, 81(5), 557–574. <https://doi.org/10.1037/0021-9010.81.5.557>
- Walton, K. E., Murano, D., Burrus, J., & Casillas, A. (2021). Multimethod support for using the Big Five framework to organize social and emotional skills. *Assessment*, 1–16. <https://doi.org/10.1177/10731911211045744>
- Zell, E., & Lesick, T. L. (2021). Big Five personality traits and performance: A quantitative synthesis of 50+ meta-analyses. *Journal of Personality*, 90, 559–573. <https://doi.org/10.1111/jopy.12683>
- Zickar, M. J., Gibby, R. E., & Robie, C. (2004). Uncovering faking samples in applicant, incumbent, and experimental data sets: An application of mixed-model item response theory. *Organizational Research Methods*, 7(2), 168–190. <https://doi.org/10.1177/1094428104263674>
- Zimmerman, R. D. (2008). Understanding the impact of personality traits on individuals' turnover decisions: A meta-analytic path model. *Personnel Psychology*, 61(2), 309–348. <https://doi.org/10.1111/j.1744-6570.2008.00115.x>



ACT[®]